

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

Обробка сигналів електронних та акустичних систем

Конспект лекцій

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня доктор філософії
за освітньо-науковою програмою
«Електроніка»,
спеціальності 171 Електроніка*

Київ
КПІ ім. Ігоря Сікорського
2021

Обробка сигналів електронних та акустичних систем: Конспект лекцій [Електронний ресурс]: навч. посіб. для здобувачів освіти рівня доктор філософії спеціальності 171 «Електроніка» / уклад.: А. М. Продеус; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 12,2 Мбайт). - Київ: КПІ ім. Ігоря Сікорського, 2021. – 249 с.

*Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 7 від 13.05.2021 р.)
за поданням Вченої ради факультету електроніки (протокол № 04/21 від 26.04.2021 р.)*

Електронне мережне навчальне видання

Обробка сигналів електронних та акустичних систем Конспект лекцій

Укладач: *Продеус Аркадій Миколайович*, д-р техн. наук, проф.

Відповідальний
редактор *Лазебний В.С.*, канд. техн. наук, доц.

Рецензент: *Шуляк О. П.*, канд. техн. наук, доц., доцент кафедри електронної інженерії КПІ ім. Ігоря Сікорського

Даний навчальний посібник складається із трьох розділів. Перший розділ містить 9 підрозділів із теоретичними відомостями, що висвітлюють такі питання як акустична паспортизація приміщень, різновиди джерел тестових сигналів, штучна голова та її застосування для акустичних вимірювань, створення та використання банку бінауральних імпульсних характеристик приміщень, методи пасивної та активної корекції акустики приміщень, програмне та апаратне забезпечення акустичної експертизи приміщень. Другий розділ навчального посібника містить 2 підрозділи, що висвітлюють питання корекції акустики приміщень. Третій розділ навчального посібника містить завдання із оцінювання розбірливості мови методом артикуляційного випробування та завдання для самостійного аналізу наукової статті оглядового характеру.

Даний навчальний посібник забезпечує проведення лекційних та практичних занять із дисципліни «Комп'ютерна обробка акустичних сигналів» для здобувачів освіти рівня підготовки PhD.

Крім того, навчальний посібник буде корисним для студентів акустичних та телекомунікаційних спеціальностей технічних навчальних закладів, що навчаються за спеціальністю «Електроніка», а також фахівцям в галузі акустичної експертизи та корекції комунікаційних каналів.

© КПІ ім. Ігоря Сікорського, 2021

ЗМІСТ

Вступ	4
Розділ 1. Обробка електричних сигналів при акустичних вимірюваннях	5
1.1. Мета та завдання навчальної дисципліни	5
1.2. Вимоги до акустики класів, аудиторій, залів	13
1.3. Акустична паспортизація аудиторій	19
1.4. Джерела тестових сигналів: переваги та недоліки	56
1.5. Штучна голова та її застосування для акустичних вимірювань	79
1.6. Банки бінауральних імпульсних характеристик	90
1.7. Акустична експертиза приміщень для мовленнєвих презентацій	108
1.8. Акустична експертиза концертних залів	132
1.9. Програмне забезпечення та апаратурне обладнання	150
Розділ 2. Методи та системи корекції акустики приміщень	165
2.2. Методи пасивної корекції акустики приміщень	165
2.3. Методи активної корекції акустики приміщень	169
Розділ 3. Завдання до МКР та самостійної роботи	228
3.1. Завдання із МКР.....	228
3.2. Завдання до самостійної роботи	228
Література	230
Базова література	230
Додаткова література	233
Додатки	237
Додаток до підрозділу 1.7. Розробка пакетів програм для суб'єктивного оцінювання розбірливості мови	237

Вступ

Мета навчальної дисципліни «Обробка сигналів електронних та акустичних систем» полягає в поглибленому вивченні теоретичних засад цифрової обробки сигналів при акустичній експертизі та корекції характеристик приміщень різного призначення – навчальних приміщень та концертних залів. Для кращого засвоєння теоретичного матеріалу планується виконання практичної роботи та самостійної роботи, що завершуються індивідуальним звітуванням у вигляді письмового звіту та участі у семінарах.

Виходячи із зазначеної мети, в навчальному посібнику висвітлюються питання акустичної паспортизації приміщень, виконується порівняльний аналіз джерел тестових сигналів, надається опис програмно-апаратного комплексу «Штучна голова» та наводяться приклади її застосування як для акустичних вимірювань, так і для створення банку бінауральних імпульсних характеристик приміщень. Крім того, надається інформація про програмне та апаратне забезпечення акустичної експертизи приміщень, методи пасивної та активної корекції акустики приміщень. Нарешті, надано завдання із оцінювання розбірливості мови методом артикуляційного випробування та завдання для самостійного аналізу наукової статті оглядового характеру.

Дана дисципліна є подальшим розвитком таких дисциплін як «Комп'ютерна обробка акустичних сигналів», «Акустичні інформаційні системи», «Акустичні вимірювання», «Акустичні прилади та системи», «Архітектурна акустика», «Психоакустика».

Навчальний посібник буде корисним також для здобувачів освіти акустичних та телекомунікаційних спеціальностей технічних навчальних закладів, що навчаються за спеціальністю «Електроніка», а також фахівцям в галузі акустичної експертизи та корекції комунікаційних каналів.

Розділ 1. Обробка електричних сигналів при акустичних вимірюваннях

Лекція 1. Мета та завдання навчальної дисципліни

Мета навчальної дисципліни «Обробка сигналів електронних та акустичних систем» полягає в поглибленому вивченні теоретичних засад цифрової обробки сигналів при акустичній експертизі та корекції характеристик приміщень різного призначення – навчальних приміщень та концертних залів. Для кращого засвоєння теоретичного матеріалу планується виконання практичної роботи та самостійної роботи, що завершуються індивідуальним звітуванням у вигляді письмового звіту та участі у семінарах.

Дана дисципліна є подальшим розвитком таких дисциплін як «Комп'ютерна обробка акустичних сигналів», «Акустичні інформаційні системи», «Акустичні вимірювання», «Акустичні прилади та системи», «Архітектурна акустика», «Психоакустика».

Розподіл навчального часу за видами занять:

- Лекції – 18 годин;
- Практичні заняття та семінари – 18 годин;
- Модульна контрольна робота – 2,25 годин;
- Консультації – 1,94 годин;
- Залік – 6 годин.
- Загалом – 46,19 годин.

Методичні особливості вивчення навчальної дисципліни полягають у засвоєнні, теоретичному та практичному, навичок наукових досліджень у вигляді генеруванні гіпотез та перевірки їх справедливості шляхом створення аналітичних, комп'ютерних та фізичних моделей об'єктів дослідження, а також шляхом експериментальних досліджень із використання слухової системи людини.

Індивідуальні семестрові завдання полягають у виконанні:

- модульної контрольної роботи (МКР) "Автоматизоване суб'єктивне оцінювання розбірливості мови в аудиторіях 209 та 438";
- самостійної роботи у вигляді опрацювання (аналізу) публікації S. Cecchi, A. Carini and S. Spors. "Room Response Equalization—A Review".

Форми звітності здобувачів освіти:

- наявність та якість конспекту оцінюється за 5-бальною шкалою;
- письмові звіти за результатами виконання МКР та самостійної роботи - оцінюються за 5-бальною шкалою;
- електронні файли із результатами виконання МКР - оцінюються за 5-бальною шкалою;
- участь в семінарі за результатами виконання самостійної роботи - оцінюються за 5-бальною шкалою.

Дистанційні ресурси:

- Дистанційний ресурс **Moodle**:
<https://do.ipk.kpi.ua/course/view.php?id=434>
- Дистанційний ресурс **Google Classroom**:
<https://classroom.google.com/u/1/c/MjU2NjA0NjU2MjQ5>

Базова література

1. Продеус А., Дидковский В., Дидковская М. Акустическая экспертиза и коррекция коммуникационных каналов. Монография. LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, Saarbrücken, Deutschland, 2017, ISBN: 978-3-330-04591-0
2. Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации. Монография. – К., «Имекс-ЛТД», 2008. – 420 с.

3. Ю.С. Костючок, Л.С. Мартинович, Д.Е. Моторнюк, В.А. Нечитайло, А.В. Храпачевский, А.Н. Продеус. Акустическая паспортизация учебных помещений. - ISSN 1811-4512. Electronics and Communications, 2016, Vol.21, No.2(91). – P.63-70
4. О. Дворник, А. Продеус, М. Дідковська, Д. Моторнюк, "Апаратно-програмний комплекс «Штучна голова». Частина 1. Коригування частотної характеристики тракту," ISSN 2523-4447, Мікросистеми, електроніка та акустика, 2020, vol. 22, no. 1, сс. 56-64. DOI: <https://doi.org/10.20535/2523-4455.mea.198431>
5. О. Дворник, А. Продеус, М. Дідковська, Д. Моторнюк, "Апаратно-програмний комплекс «Штучна голова». Частина 2. Оцінювання розбірливості мови в аудиторіях," Microsystems, Electronics and Acoustics, vol. 22, no. 3, 2020, pp. 48-55. DOI: <https://doi.org/10.20535/2523-4455.mea.209928>
6. Продеус А.М. Комп'ютерна обробка акустичних сигналів [Електронний ресурс]. Навчальний посібник. Київ: КПІ, 2018, 237 с. <https://ela.kpi.ua/handle/123456789/25161>
7. Prodeus, A., Bukhta, K., Morozko, P., Serhiienko, O., Kotvytskyi, I., Shherbenko I. (2018). Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility. IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO`2018), Kyiv, Ukraine, 533-538. doi: <https://doi.org/10.1109/ELNANO.2018.8477568>
8. Сергиенко А.Б. Цифровая обработка сигналов. СПб: Питер., 2002 г., 606 с.
9. Богнер Р., Константи́нидис А. (ред.) Введение в цифровую фильтрацию. М.: Мир, 1976, 216 с.
10. Leccese, F., Rocca, M., Salvadori, G. (2018) Fast estimation of Speech Transmission Index using the Reverberation Time: Comparison between predictive equations for educational rooms of different sizes. Applied Acoustics, 140, 143–149. doi: <https://doi.org/10.1016/j.apacoust.2018.05.019>

11. Nocke, C. (2016). New standards in Architectural Acoustics - a German view. ACOUSTICS`2016, Brisbane, Australia. Available in: https://www.acoustics.asn.au/conference_proceedings/AASNZ2016/papers/p114.pdf
12. DIN 18041:2004-05. (2004). Hörsamkeit von kleinen und mittleren Räumen. (Acoustical quality in small to medium-sized rooms.),” Beuth Verlag GmbH. Available in: <https://www.beuth.de/de/norm/din-18041/69640391>
13. ANSI/ASA S12.60. (2015). Acoustical performance criteria, design requirements, and guidelines for schools. Part 1: permanent schools. American National Standards Institute, Inc. (ANSI) and Acoustical Society of America (ASA) New York, USA. Available in: <https://webstore.ansi.org/standards/asa/ansiasas12602010partr2015>
14. International standard ISO 3382-1. (2009). Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces. First edition. Available at: <https://www.iso.org/obp/ui/#iso:std:iso:3382:-1:ed-1:v1:en>
15. British standard BS EN 60268-16:2011. (2012). Sound systems equipment - Part 16: Objective rating of speech intelligibility by speech transmission index. BSI. Available at: https://infostore.saiglobal.com/en-us/Standards/BS-EN-60268-16-2011-217511_SAIG_BSI_BSI_512663/
16. Acoustic design of schools: performance standards. (2015). Building bulletin 93. UK Department for Education, UK Education Funding Agency. Available at: <https://www.bloconconsulting.co.uk/building-bulletin-93-bb93/>
17. Canning, D., James, A. (2012). The Essex Study. Optimised classroom acoustics for all. The Association of Noise Consultants. Available at: https://www.ndcs.org.uk/media/1808/the_essex_study_optimised_classroom_acoustics_for_all1.pdf
18. IEC 60268-16. (2016). Sound system equipment – Part 16: objective rating of speech intelligibility by speech transmission index. International Electrotechnical Commission (IEC), Geneva, Switzerland. Available at:

Додаткова література

19. Дідковський В.С., Дідковська М.В., Продеус А.М. Комп'ютерна обробка акустичних сигналів. Навчальний посібник. – К., «Імекс-ЛТД», 2010. – 420 с.
20. Стандарти та вимоги до акустики шкільних класів та аудиторій вишів Америки, Великобританії, Німеччини, Росії, України. Доступно в системі дистанційного навчання НТУУ «КПІ ім. І. Сікорського» та в Google Classroom.
21. M. Jeub, M. Schafer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in Proc. Int. Conference on Digital Signal Processing (DSP), Santorini, Greece, 2009.
22. Eggenschwiler, K. (2005). Lecture Halls - Room Acoustics and Sound Reinforcement. Forum Acusticum, Budapest, 1-7. Available at: https://www.researchgate.net/publication/238114069_Lecture_Halls_-_Room_Acoustics_and_Sound_Reinforcement
23. Bradley, J., Reich, R., Norcross, S. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. Journal of the Acoustical Society of America, 106 (4), Pt. 1. doi: <https://doi.org/10.1121/1.427932>
24. Bradley, J., Sato, H., Picard, M. (2003). On the importance of early reflections for speech in rooms. Journal of the Acoustical Society of America, 113 (6), 3233-3244. doi: <http://dx.doi.org/10.1121/1.1570439>
25. Hu, Y., Kokkinakis, K. (2013). Effects of early and late reflections on intelligibility of reverberated speech by cochlear implant listeners. Journal of the Acoustical Society of America, 135 (1). doi: <https://doi.org/10.1121/1.4834455>

26. Yang, W., Bradley, J. (2009). Effects of room acoustics on the intelligibility of speech in classrooms. *Journal of the Acoustical Society of America*, 125 (2), 1-12. doi: <https://doi.org/10.1121/1.3058900>
27. Bradley, J. (2011). Review of objective room acoustics measures and future needs. *Applied Acoustics*, 72 (10), 713-720. doi: <https://doi.org/10.1016/j.apacoust.2011.04.004>
28. Arweiler, I., Buchholz, J., Dau T. (2009). Speech intelligibility enhancement by early reflections. *ISAAR 2009: Binaural Processing and Spatial Hearing*, 2nd International Symposium on Auditory and Audiological Research, Elsinore, Denmark. Available at: <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-29>
29. Sari, L., Utami S., Sarwono, J. (2018). Objective and subjective acoustics measurement of audience seating areas in a medium size auditorium. *Journal of Physics: Conf. Series* 1075. doi: 10.1088/1742-6596/1075/1/012015
30. Choi, Y.-J. (2020). The intelligibility of speech in university classrooms during lectures. *Applied Acoustics*, 162, 107211. doi: <https://doi.org/10.1016/j.apacoust.2020.107211>
31. Eldakdoky, S. (2017). Optimizing acoustic conditions for two lecture rooms in Faculty of Agriculture, Cairo University. *Ain Shams Engineering Journal*, 8, 481–490. doi: <http://dx.doi.org/10.1016/j.asej.2016.08.013>
32. Prodeus, A., Didkovska, M., Motorniuk, D., Dvornyk, O. (2020). The Effects of Noise, Early and Late Reflections on Speech Intelligibility. *IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO`2020)*, Kyiv, Ukraine, 488–492. doi: <https://doi.org/10.1109/ELNANO50318.2020.9088854>
33. Prodeus, A., Didkovska, M. (2020). Objective assessment of speech intelligibility in small and medium-sized classrooms. *IEEE Int. Scientific-Practical Conf. on Problems of Infocommunications, Science and Technology (PIC S&T`2020)*, Kharkiv, Ukraine. Available at: https://www.researchgate.net/publication/347490796_Objective_Assessment_of_Speech_Intelligibility_in_Small_and_Medium-Sized_Classrooms

34. Prodeus, A., Didkovska, M., Kukharicheva, K., Motorniuk, D. (2020). Modeling the Influence of Early Sound Reflections on Speech Intelligibility. IEEE 6th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC`2020), Kyiv, Ukraine, 47-50. Available at: https://www.researchgate.net/publication/346080603_Modeling_the_Influence_of_Early_Sound_Reflections_on_Speech_Intelligibility
35. Jacob, K. (1988). Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms. 85th Convention of the Audio Engineering Society, Los Angeles, 1020-1030. Available at: <https://secure.aes.org/forum/pubs/journal/?elib=6056>
36. Lochner J., Burger, J. (1964). The influence of reflections on auditorium acoustics,” J. Sound Vib. 1, 426–454. doi: [https://doi.org/10.1016/0022-460X\(64\)90057-4](https://doi.org/10.1016/0022-460X(64)90057-4)
37. Bistafa, S., Bradley, J. (2000). Reverberation time and maximum background noise level for classrooms from a comparative study of speech intelligibility metrics. J. Acoust. Soc. Am., 107, 861–875. doi: <https://doi.org/10.1121/1.428268>
38. Gold, B., Morgan, N., Ellis, D. (2011). Speech and audio signal processing. Processing and Perception of Speech and Music. John Wiley & Sons, Inc., 679. Available at: <https://ieeexplore.ieee.org/book/8039862>
39. Nestoras, C., Dance, S. (2013). The Interrelationship Between Room Acoustics Parameters as Measured in University Classrooms Using Four Source Configurations. Building acoustics, 20 (1), 43–54. doi: <https://doi.org/10.1260/1351-010X.20.1.43>
40. Aachen Impulse Response Database. Available at: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>
41. Duran, S., Ausiello, L., Battaner-Moro, J. (2019). Acoustic Design Criteria for Higher-Education Learning Environments. Proceedings of the Institute of Acoustics, 41 (3), 1-12. Available at:

<https://pure.solent.ac.uk/en/publications/acoustic-design-criteria-for-higher-education-learning-environment>

42. Jedidi, M., Boulila, A. (2016). Acoustic study of an auditorium by the determination of reverberation time and speech transmission index. *Int. J. Architect. Eng. Urban Plan*, 26(1), 25-32. Available at: <http://ijaup.iust.ac.ir/article-1-312-en.pdf>
43. Silva, L., Santos, R. (2013). Acoustical Comfort in Primary School Classrooms in the City of Joao Pessoa, Paraiba, Brazil. *Journal of Ergonomics*, S:1. doi: <https://doi.org/10.4172/2165-7556.S1-001>
44. Rennies, J., Schepker, H., Holube, I., Kollmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J. Acoust. Soc. Am.*, 136 (5), 2642–2653. doi: <https://doi.org/10.1121/1.4897398>
45. Kuttruff, H. (2009). *Room Acoustics*. Fifth edition. Spon Press 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN. Available at: <https://www.amazon.com/Room-Acoustics-Fifth-Heinrich-Kuttruff/dp/0415480213>
46. Steeneken, H. (2014). Forty years of speech intelligibility assessment (and some history). *Proceedings of the Institute of Acoustics*, 36, Pt.3. Available at: <http://www.steeneken.nl/wp-content/uploads/2014/04/IOA-Forty-years-keynote.pdf>

Лекція 1.2. Вимоги до акустики класів, аудиторій, залів

Однією з умов отримання якісної освіти є хороша розбірливість мови викладача в місцях розташування слухачів. Тим часом, непродумані архітектурні рішення часто призводять до незадовільної розбірливості в нових або реконструйованих аудиторіях. Цим пояснюється увага дослідників різних країн до проблеми забезпечення хорошої акустики в навчальних приміщеннях. Важливим наслідком таких досліджень є обґрунтування рекомендацій і вимог до акустики шкільних приміщень.

Одним із важливих результатів досліджень акустики шкільних класів є кількісна оцінка залежності розбірливості мови, спотвореної шумом або реверберацією, від віку слухачів. Іншим важливим результатом є кількісна оцінка ступеня чутливості слухачів з втратами слуху до впливу шуму і реверберації.

Хоча шум є значно більш деструктивним фактором, в порівнянні з реверберацією, велику кількість робіт присвячено саме аналізу впливу реверберації на розбірливість мови. Пояснити це можна тим, що ранні відображення звуку, на відміну від пізніх відображень, підвищують відношення сигнал-шум, що спричиняє підвищення розбірливості мови.

При бінауральному прослуховуванні мовних сигналів відношення сигнал-шум підвищується на 1-2 дБ, що призводить до додаткового підвищення розбірливості мови, в порівнянні з монауральним прослуховуванням. У зв'язку з цим важливим завданням стає створення банків бінауральних імпульсних характеристик (БІХ) різних приміщень і, в тому числі, університетських аудиторій.

Якщо розбірливості мови в шкільних класах приділялося багато уваги, не таким є стан справ із аудиторіями університетів, які істотно відрізняються від шкільних класів розмірами, формою, оздоблювальними матеріалами.

Таким чином, актуальність проблеми оцінювання розбірливості мови в університетських аудиторіях зумовлена, щонайменше, двома факторами. З

одного боку, це недостатня, до недавнього часу, увага до акустичних властивостей університетських аудиторій і, як наслідок, відсутність чітких рекомендацій щодо вибору простих, але ефективних мір розбірливості мови. З іншого боку, очевидна потреба в таких рекомендаціях з боку адміністрації університетів, а також архітекторів і дизайнерів, що проектують або реконструюють університетські приміщення.

Акустика шкільних приміщень досить є досить добре вивченою [1-8]. Приклади невдалих архітектурних рішень, що призводять до зниження розбірливості мови через підвищення рівня шуму і часу реверберації, появи відлуння або пурхаючого відлуння (flutter echo), наведені в [1]. Як показали експериментальні дослідження, шум у класних приміщеннях помітніше погіршує розбірливість мови, ніж пізні відбиття звуку [2]. Завдяки раннім відбиттям звуку досягається прийнятна розбірливість мови в приміщеннях, навіть якщо голова диктора спрямована в протилежну від слухача сторону [3]. Проблема забезпечення гарної розбірливості мови є особливо актуальною для школярів з вадами слуху [4] і школярів молодших класів [5], для яких рівень шуму не повинен перевищувати 30-35 дБА. Незважаючи на значний обсяг проведених досліджень, в [6] зроблено висновок про існування ряду питань, які потребують подальшого розвитку. Так, наприклад, наведені в [7] результати не підтвердили висновок роботи [5] про еквівалентність дії ранніх віддзеркалень і прямого звуку. У розвиток представлених в [5] і [7] результатів, в [8] зроблено припущення, що процес інтегрування ранніх відбиттів з прямим звуком відбувається незалежно від просторової обробки звуку слуховою системою людини.

Про досить хорошу вивченість акустики шкільних приміщень також свідчить наявність низки національних і міжнародних стандартів і рекомендацій. Наприклад, в стандарті DIN 18041: 2004-05 (Німеччина) приміщення розділені на категорії А і В, й для приміщень категорії А («музика», «мова», «спорт», «спілкування») вказано оптимальні значення часу реверберації. Вимоги та рекомендації до акустичних характеристик і шумоізоляції шкільних приміщень

викладені в стандарті ANSI/ASA S12.60 / Part 1 (США). Прямі та непрямі методи оцінювання індексу розбірливості мови STI і його модифікацій (STIPA, STITEL, RASTI) описані в британському стандарті BS EN 60268-16: 2011 і міжнародному стандарті IEC 60268-16. Рекомендації з врахування потреб звичайних школярів і учнів з вадами слуху, викладені у збірнику «Acoustic design of schools: performance standards. Building bulletin 93» (Великобританія). Аналогічні рекомендації, підкріплені результатами експериментальних досліджень, викладені у збірнику «The Essex Study. Optimised classroom acoustics for all» (Великобританія).

Також досить добре вивчена акустика концертних залів. Перелік найбільш важливих акустичних характеристик концертних залів, а також алгоритми їх оцінювання представлено в міжнародному стандарті ISO 3382-1. Що стосується акустики університетських аудиторій, в явному вигляді відповідні рекомендації практично відсутні. Можливою причиною такої ситуації може бути «проміжне» становище університетських аудиторій.

Дійсно, аналіз характеристик типових університетських аудиторій свідчить, що середній розмір університетських аудиторій більше за такий для шкільних класів, але менше середнього розміру концертних залів. Наприклад, в [9] до аудиторій середнього розміру віднесено приміщення з балконом, яке розраховане на 1100 слухачів та має форму «віяла». Об'єм розглянутих в [10] дванадцяти університетських аудиторій малого і середнього розміру, розрахованих на 30-70 студентів, варіюється від 188 м³ до 343 м³. Акустичні характеристики аудиторій прямокутної форми, об'ємом 530 м³ та 1740 м³, розглянуті в [11]. В [12] зроблена спроба класифікувати аудиторії за розміром. При цьому виділено три групи приміщень малого, середнього і великого розмірів, об'ємом менше 350 м³, 350-650 м³ та більше 650 м³, відповідно. Аналогічну класифікацію приміщень запропоновано в [13], де до малих, середніх і великих аудиторій віднесено приміщення об'ємом менше 230 м³, 230-350 м³ і більше 350 м³, відповідно. Типовою особливістю великих аудиторій є похила підлога [11], [12].

Разом з тим, університетські аудиторії є меншими за концертні приміщення. Наприклад, об'єм розглянутих в [14] десяти залів Бостона (США), призначених для мовленнєвих та музичних презентацій, варіюється від 3000 до 60000 м³. Оздоблення університетських аудиторій мало відрізняється від такого в шкільних класах, проте воно, як правило, є менш різноманітним, ніж в концертних залах [14].

З огляду на недостатню вивченість особливостей акустики університетських аудиторій, виглядає зрозумілим зростання інтересу дослідників до цієї проблеми, що сприяє отриманню низки корисних рекомендацій. Наприклад, оцінювання розбірливості мови (індекс STI) в дванадцяти університетських аудиторіях, обладнаних системами звукопідсилення, дозволило зробити висновок, що такі системи потрібні лише за умов значного рівня фонового шуму [1]. Процедура оцінювання розбірливості мови суб'єктивним (артикуляційним) методом значно спростилася. Так, замість описаної в [3] і розвиненої в [7] громіздкої та дорогої організації тестування, тестові сигнали стали частіше синтезувати шляхом згортки чистого сигналу з БІХ приміщення [9]. На жаль, більшість публікацій, присвячених аналізу акустичних характеристик університетських аудиторій, носить фрагментарний характер.

Зокрема, аналітичний опис характеру і ступеня впливу відбиваючих поверхонь на рівень звуку поблизу цих поверхонь було виконано досить давно. Проте, в літературі приділено недостатню увагу кількісній оцінці ступеня цього впливу на розбірливість мови. Окремі отримані результати такого роду носять скоріше випадковий, ніж систематичний, характер. Тому актуальним є питання заповнення зазначених прогалів.

Література до розділу 1.2

[1] Eggenschwiler, K. (2005). Lecture Halls - Room Acoustics and Sound Reinforcement. Forum Acusticum, Budapest, 1-7. Available at:

[https://www.researchgate.net/publication/238114069_Lecture_Halls_-
_Room_Acoustics_and_Sound_Reinforcement](https://www.researchgate.net/publication/238114069_Lecture_Halls_-_Room_Acoustics_and_Sound_Reinforcement)

[2] Bradley, J., Reich, R., Norcross, S. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. *Journal of the Acoustical Society of America*, 106 (4), Pt. 1. doi: <https://doi.org/10.1121/1.427932>

[3] Bradley, J., Sato, H., Picard, M. (2003). On the importance of early reflections for speech in rooms. *Journal of the Acoustical Society of America*, 113 (6), 3233-3244. doi: <http://dx.doi.org/10.1121/1.1570439>

[4] Hu, Y., Kokkinakis, K. (2013). Effects of early and late reflections on intelligibility of reverberated speech by cochlear implant listeners. *Journal of the Acoustical Society of America*, 135 (1). doi: <https://doi.org/10.1121/1.4834455>

[5] Yang, W., Bradley, J. (2009). Effects of room acoustics on the intelligibility of speech in classrooms. *Journal of the Acoustical Society of America*, 125 (2), 1-12. doi: <https://doi.org/10.1121/1.3058900>

[6] Bradley, J. (2011). Review of objective room acoustics measures and future needs. *Applied Acoustics*, 72 (10), 713-720. doi: <https://doi.org/10.1016/j.apacoust.2011.04.004>

[7] Arweiler, I., Buchholz, J., Dau T. (2009). Speech intelligibility enhancement by early reflections. *ISAAR 2009: Binaural Processing and Spatial Hearing*, 2nd International Symposium on Auditory and Audiological Research, Elsinore, Denmark. Available at: <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-29>

[8] Warzybok, A., Rennies, J., Doclo, S., Kollmeier, B. (2011). Influence of early reflections on speech intelligibility under different noise conditions. *Forum Acusticum*, Aalborg, Denmark. Available at: https://uol.de/f/6/dept/mediphsik/ag/sigproc/download/papers/SP2011_5.pdf

[9] Sari, L., Utami S., Sarwono, J. (2018). Objective and subjective acoustics measurement of audience seating areas in a medium size auditorium. *Journal of Physics: Conf. Series* 1075. doi: 10.1088/1742-6596/1075/1/012015

[10] Choi, Y.-J. (2020). The intelligibility of speech in university classrooms during lectures. *Applied Acoustics*, 162, 107211. doi: <https://doi.org/10.1016/j.apacoust.2020.107211>

[11] Eldakdoky, S. (2017). Optimizing acoustic conditions for two lecture rooms in Faculty of Agriculture, Cairo University. *Ain Shams Engineering Journal*, 8, 481–490. doi: <http://dx.doi.org/10.1016/j.asej.2016.08.013>

[12] Leccese, F., Rocca, M., Salvadori, G. (2018) Fast estimation of Speech Transmission Index using the Reverberation Time: Comparison between predictive equations for educational rooms of different sizes. *Applied Acoustics*, 140, 143–149. doi: <https://doi.org/10.1016/j.apacoust.2018.05.019>

[13] Nestoras, C., Dance, S. (2013). The Interrelationship Between Room Acoustics Parameters as Measured in University Classrooms Using Four Source Configurations. *Building acoustics*, 20 (1), 43–54. doi: <https://doi.org/10.1260/1351-010X.20.1.43>

[14] Jacob, K. (1988). Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms. 85th Convention of the Audio Engineering Society, Los Angeles, 1020-1030. Available at: <https://secure.aes.org/forum/pubs/journal/?elib=6056>

Лекція 1.3. Акустична паспортизація аудиторій

1.3.1. Практичні засади акустичної паспортизації приміщень

Перш ніж розглядати теоретичні засади акустичної паспортизації приміщень, корисно проаналізувати, як поняття «акустичний паспорт» та «акустична паспортизація» трактують сучасні комерційні компанії України та Росії.

1.3.1.1. Акустичне проектування, компанія Tecsound (Україна) [3]

Акустичне проектування - це створення «правильної акустики» для комфорту. Основа акустичного проектування - оптимізація часу реверберації, усунення ефекту «пурхаюче відлуння». Акустичний проект виконується згідно з нормативами ДБН В.1.1-31 - 35 та галузевими нормами.

Акустичний проект містить в собі:

- визначення джерел шуму і їх шумових характеристик;
- вимірювання акустичних і фізичних характеристик приміщення;
- розрахунок очікуваних рівнів шуму;
- визначення необхідного зниження рівнів шуму;
- комп'ютерне моделювання приміщення і поширення в ньому звуку;
- розробка комплексу будівельно-акустичних дій для забезпечення нормативного шумового режиму;
- очікувані результати - ефективність проведених заходів;
- вимірювання після проведення акустичних заходів щодо захисту від шуму;
- видача **акустичного паспорту** приміщення.

Акустичний проект розв'язує такі задачі:

- звукоізоляція (стін, перекрить, перегородок, вентиляційних та інших технічних систем);
- звукопоглинання (акустика всередині приміщення);

- віброізоляція (будівель, технічного обладнання);
- акустичний дизайн;
- акустичний аудит.

Акустичний аудит - це перевірка будівельних проектів на їх відповідність чинному в Україні законодавству в сфері захисту від шуму. Дешевше і швидше виправити неточності в проекті на ранньому етапі. Можливо, невелика зміна в проекті дозволить заощадити на подальших переробках і використання екзотичних систем. Це визначення ще на етапі проектування можливих проблем зі звукоізоляцією, віброізоляцією, високими рівнями шуму в приміщеннях і на територіях, а також шляхи вирішення цих проблем.

Акустичний аудит включає в себе такі заходи:

- вимірювання;
- комп'ютерне моделювання.

1.3.1.2. Акустика приміщень, компанія Acoustic Traffic LLC (Україна) [5]:

Акустичний дизайн нерозривно пов'язаний з архітектурним дизайном приміщення. Одним із основних завдань інженера-акустика є підбір таких акустичних матеріалів, які б повністю вирішували завдання забезпечення високого акустичного комфорту в проектованому приміщенні і при цьому їх зовнішній вигляд не суперечив би концепції дизайну, запропонованої архітектором.

1.3.1.3. Компанія Security Media Systems PRO (Росія) [1]

Для підготовки аудіо-проекту перше, що необхідно зробити, це підготувати «Акустичний паспорт» шляхом моделювання, розрахунків, вимірювання:

- рівня звукового тиску;
- рівня прямого звуку;

- рівня відбитого звуку;
- відношення рівнів прямого і відбитого звуку;
- нерівномірність поля рівнів;
- нерівномірність АЧХ;
- шумові криві за результатами оцінки NC;
- показники розбірливості мови за шкалами: Alcons%, STI, RASTI;
- показники чіткості мови за шкалами: C80, C50, C20;
- ауралізація звуку з отриманням звукового файлу при різних показниках;
- висновки і рекомендації (архітектурно-акустична адаптація);
- комплексне оцінювання результатів архітектурно-акустичної адаптації;
- електроакустичні вимірювання якості роботи аудіосистеми;
- власне **«Акустичний паспорт»** у вигляді зведення загальних результатів і показників працюючої системи.

1.3.1.2. Акустична експертиза, компанія ASPD “Архітектура та Акустика” (Росія) [2]

За результатами акустичної експертизи складається **акустичний паспорт** приміщення, який може включати наступні параметри:

- спектральний аналіз рівнів шумів в приміщенні;
- рівень ізоляції ударного шуму міжповерховими перекриттями;
- рівень ізоляції повітряного шуму огорожувальними конструкціями;
- час реверберації RT60;
- імпульсні характеристики приміщення;
- розрахунок індексів чіткості C50, прозорості C80, просторового враження R, індексу розбірливості мови RASTI.

Акустичні вимірювання виконуються відповідно до ГОСТ 27296-87 (СТ. СЕВ 4866-84), при їх проведенні використовується таке ліцензійне обладнання:

- точно відкалібрований мікрофон та передпідсилювач фірми IBF-Akustik. (Ingenieurbuero f. Akustik, Elektronik, Messtechnik Alexanderstr 66 DE-68519 Viernheim Germany);
- прецизійний шумомір Testo 816;
- програмне забезпечення ETF фірми Acoustic Soft Inc. (260 Wellesley St E PH14 Toronto Ontario Canada M4X 1G6);
- програмне забезпечення Sample Champion фірми Purebits (Author Paolo Guidorzi, Release Date 15 02 2004);
- програмне забезпечення "Wave Lab" і "Sound Forge Pro 10.0."

1.3.1.4. Акустичний проект, ООО «АЭРФОН» (Росія) [4]:

Акустичне проектування - створення сприятливого середовища для прослуховування музики за допомогою застосування спеціалізованих акустичних матеріалів, конструкцій підлог, стін, стель, різних модулів, з певними акустичними властивостями, що сприяють повному розкриттю всіх особливостей звукових систем в кімнаті прослуховування. Сюди ж входить розрахунок звукоізоляції і віброізоляції підлоги, стін, стель для мінімізації впливу зовнішнього шуму і відсічення проходу звукової хвилі в інші приміщення.

Одна з головних задач акустичного проектування - оптимізація часу реверберації. Важливо також усунути ефект "пурхаючого відлуння".

Акустичне проектування може бути самостійним розділом проектної документації, або частиною технологічного розділу Технічного і Дизайн-проектів. Розраховуються кімнати, до яких пред'являються особливі технологічні вимоги до поширення звукових хвиль, їх властивостей із звукопоглинання, звукоізоляції (домашні кінотеатри, музичні кімнати прослуховування, домашні студії та ін.). У складі акустичного проектування проводиться розрахунок параметрів акустичних властивостей кімнат, відповідно до їх призначення.

Акустичний проект представляється у вигляді робочих креслень, із вказанням деталей щодо виготовлення, влаштування стін, підлоги, стелі, конструктивів, в тому числі, із застосуванням загально-будівельних матеріалів (мінеральна вата, гіпсокартон, фанера, ДСП та ін.), спеціалізованих матеріалів для акустичної корекції приміщення, виконаних на основі або з урахуванням дизайн-проекту. Проект включає точні специфікації застосовуваних на об'єкті оздоблювальних, звукопоглинальних і звукоізоляційних матеріалів.

Етапи створення проекту:

- формування технічного завдання;
- створення технічного рішення з архітектури та акустики приміщення;
- опрацювання деталей технічних завдань, що розв'язуються, коригування Технічного і Дизайн-проектів з урахуванням вимог акустичного проектування;
- створення робочої технічної документації, що є керівництвом для монтажників, будівельників з урахуванням всіх існуючих будівельних норм і вимог.

Пояснювальна записка акустичного проекту:

А. Пояснювальна записка:

- Вступ
- Визначення розмірів, форми приміщення
 - розрахунок об'єму, середніх розмірів приміщення
 - побудова профіля підлоги, стін, стелі
- Акустичний розрахунок приміщення
 - вибір звукопоглинальних матеріалів, конструкцій
 - визначення розташування, площ звукопоглинальних матеріалів
 - визначення оптимального часу реверберації
 - корегування фонду звукопоглинання приміщення
- Визначення енергетичних параметрів акустики приміщення

- визначення інтенсивності, затухання звукової енергії
- розрахунок та побудова структури ранніх відбить
- розрахунок спектра власних НЧ частот приміщення
- розрахунок індексів чіткості, музичної ясності, просторового враження
- Розрахунок рівнів шумів в приміщенні
 - визначення зовнішніх джерел шумів, їх частотний спектр
 - розрахунок звукоізоляції приміщення кінозалу
 - розрахунок результуючої частотної характеристики рівній шумів
 - Перелік компенсаційних заходів
- Результуюча акустична модель приміщення
 - комп'ютерне моделювання приміщення
 - характеристики звуку в умовах розрахованого приміщення
- Література

Б. Креслення:

- побудова профілю підлоги, стін, стелі приміщення;
- розгортка приміщення із зображенням місць установки звукопоглинальних матеріалів;
- фонд звукопоглинання приміщення;
- розріз, плани приміщення із зображенням профілювання його площ;
- картина ранніх віддзеркалень в зоні прослуховування;
- схеми виконання конструкцій в приміщенні з урахуванням звукоізоляції;
- графіки розрахункового / оптимального часу реверберації;
- специфікація звукопоглинальних і звукоізоляційних матеріалів;
- **акустичний паспорт приміщення.**

П'ятий, заключний, етап архітектурного проектування - проводиться втілення акустичного проекту, в узгодженій послідовності з Технічним, Дизайн-проектами.

1.3.2. Теоретичні засади акустичної паспортизації приміщень (огляд літературних джерел)

1.3.2.1. Алдошина, Пріттс. Акустика лекційних залів (аудиторій) [6]:

Акустичне проектування приміщень, призначених в основному для мовленнєвих програм, повинно виходити з виконання наступних завдань:

- забезпечення максимальної *розбірливості* мови, для того щоб донести до слухачів семантичну (сміслову) інформацію;
- збереження природного *тембру* голосу і його змін для передачі естетичної інформації;
- забезпечення правильної *локалізації* звукового образу, поєднання його з глядацьким.

Основні фактори, що визначають високий відсоток *розбірливості*, такі:

- високий рівень прямого звуку на всіх слухацьких місцях і максимальне відношення рівня прямого звуку до пізніх відбить, які прийшли до слухачів після 80 мс;
- порівняно невеликий, від 0,4 с до 1,2 с, час стандартної реверберації;
- досить високий рівень перших дискретних відбить, які надходять до слухачів протягом перших 80 мс;
- високий рівень відношення сигнал-шум;
- відсутність луни та інших завад в приміщенні.

Збереження тембральних характеристик мови визначається в першу чергу відсутністю резонансів, що особливо актуально для приміщень малих розмірів, оскільки резонанси можуть посилювати окремі обертони в спектрі голосу і приводити до його «фарбування» (особливо це ймовірно для чоловічих голосів в діапазоні 80-300 Гц).

Порушення правильної локалізації в приміщеннях з «природною акустикою» може бути викликано наявністю сильних пізніх відбить (луни або пурхаючої луни), для усунення яких потрібні спеціальні конструктивні заходи.

1.3.2.2. Eggenschwiler K. Lecture Halls - Room Acoustics and Sound Reinforcement [7].

Стаття написана швейцарським дослідником і представляє результати дослідження великих лекційних аудиторій, обладнаних системами звукопідсилення та індукційними петлями на допомогу людям із вадами слуху.

Вимоги до шумності приміщень для людей із вадами слуху представлені в таблиці:

Requirements	Maximum Noise Level
low	40 dB(A)
middle	35 dB(A)
high	30 dB(A)

Table 1: Maximum allowable background noise levels in accordance with DIN 18041

Вимоги до часу реверберації для заповнених приміщень представлені на Fig. 2. Для незаповнених приміщень час реверберації може бути вищим не більше ніж на 0,2 с:

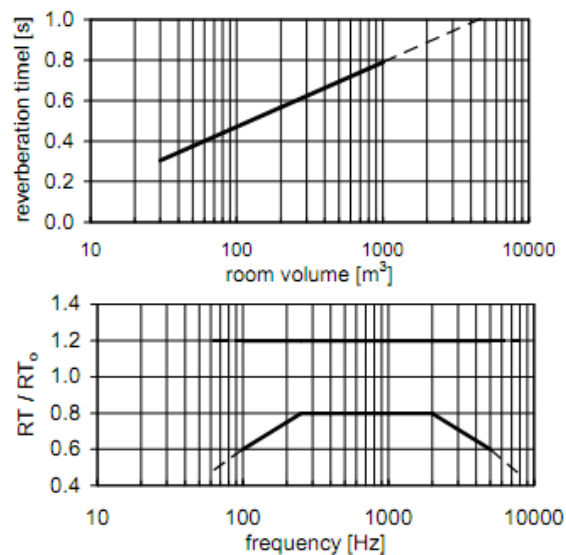


Figure 2: Optimum reverberation time for occupied instruction rooms for the octave bands 500 Hz and 1000 Hz (above) and band of tolerance for the recommended reverberation time as a function of frequency (below) according to DIN 18041. RT_0 = optimal reverberation time according to the upper diagramm, RT = reverberation time.

Вимоги до форми приміщень досить добре відомі, вони викладені в стандартах і технічній літературі. На Fig. 3 показано увігнуту задню стінку, що створює пурхаюче відлуння.



Figure 3: Rear wall of a lecture hall (not the same room as in figure 1), which throws an annoying echo back toward the front area, with or without the sound system.

Мінімальні *вимоги до розбірливості* мовлення:

Common Intelligibility Scale CIS	≥ 0.75
Speech Transmission Index STI	≥ 0.56
Definition	≥ 0 dB
Articulation Loss of Consonants ALcons	< 8 %

Вимоги до рівня звуку на місцях слухачів - при несуттєвому рівні фонового шуму рівень звуку повинен становити 65-75 дБ 6 дБ.

Вимоги до локалізації і природності тембру не є головними, проте вони важливі. Для гарної локалізації слід правильно відрегулювати часову затримку. Для природності тембру слід врахувати, що система звукопідсилення не повинна підкреслювати низькі частоти.

Вимоги до мікрофонів - доцільно використовувати гарнітури, що забезпечують однакову дистанцію від мікрофона до рота.

Допомога людям із вадами слуху:

- використання систем з індукційними петлями;
- використання систем з частотною модуляцією;
- використання систем з інфрачервоними датчиками.

Типові недоліки систем озвучування:

- гучномовці невдало встановлені і орієнтовані;
- частотні характеристики оптимізовані під музику, але не під мову;
- системи з кількох гучномовців недостатньо узгоджені;
- мікрофонні входи не забезпечують відсутність перевантаження, а користуватися ними надто складно;
- багато систем важко використовувати недосвідченим користувачам;
- системи допомоги людям із вадами слуху відсутні або працюють погано;
- через погану акустику приміщень доводиться використовувати системи придушення зворотного акустичного зв'язку;
- проблеми через заборонені зони розміщення гучномовців, спричинені розміщенням комп'ютерних проекторів;

- перевірені гучномовці відсутні в продажу.

1.3.2.3. Youssef R. Acoustical Quality Assessment of Lecture halls at Lund University, Sweden [8]:

Проводилася оцінка акустики лекційних приміщень в Лундському університеті (Швеція), заснованому в 1666 р. Оцінювалися як старі, так і недавно побудовані приміщення. Проблемою є все більш широке використання технічних аудіовізуальних засобів. При цьому аудіовізуальне обладнання встановлюється в приміщеннях, не розрахованих для цього. Як наслідок, засвоєння інформації студентами може погіршитися.

Мета роботи полягала в зіставленні кількох приміщень університету за такими показниками як час реверберації, сила звуку, чіткість і розбірливість мови відповідно до стандарту ISO 3382-1.

Розглянуто три приміщення (Figs. 1-3, Table 1).



Figure 1 – Classroom P1



Figure 2 – Classroom N1



Figure 3 – Classroom 'Navet'

Table 1 – Data for the measured classrooms

Classrooms	Volume, m ³	Seats	Volume/Seat m ³
P1	169	25	6.76
N1	220	36	6
Navet	575

Перші два приміщення - це звичайні невеликі аудиторії. Третє приміщення дещо незвичне - воно порожнє, об'ємне, використовується для різних заходів, таких як конференції, курсові заняття, презентації. Його стіни не паралельні, висота стелі 3,75 м, воно містить 3 скляних двері на одній стіні і дві таких же двері - на іншій стіні.

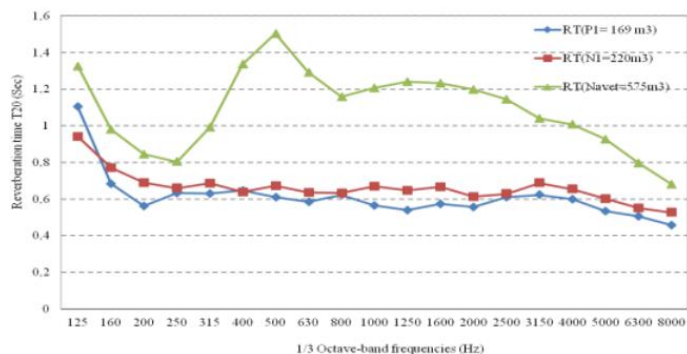
Організація експерименту

Джерела звуку та мікрофони розміщувалися на висоті 1,5 м від підлоги на місцях, типових для викладача і студентів. Вікна та двері були закриті, кондиціонери включені, вимірювання проводилися в денний час. На джерело звуку подавався білий шум, запис звуку з виходів мікрофонів проводився протягом 10 секунд. Джерело звуку розміщувався поперемінно в 2-х позиціях, 5 мікрофонів в кожному випадку розміщувалися заново на найбільш популярних місцях.

Результати експериментів

Оцінювалися і порівнювалися з рекомендованими значеннями наступні параметри: час реверберації T20, сила звуку G, ясність C50 і розбірливість мови STI.

Результати оцінювання T20 представлені на рисунку:



Видно, що час реверберації зростає зі збільшенням об'єму приміщення. При цьому найбільш цікавим є, з точки зору умов слухання в класах, середній діапазон частот 500-2000 Гц. Час реверберації в перших двох приміщеннях більш сприятливий для слухання, це пояснюється більш сильним поглинанням звуку.

Крім того, отримано графіки розподілу по площі приміщення розрахункового і реального звукового тиску.

Залежність ясності мови від частоти показана на малюнку:

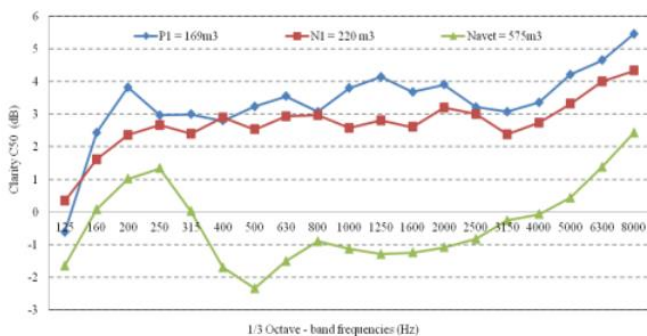


Figure 7 – Clarity C50 against frequency for all classrooms

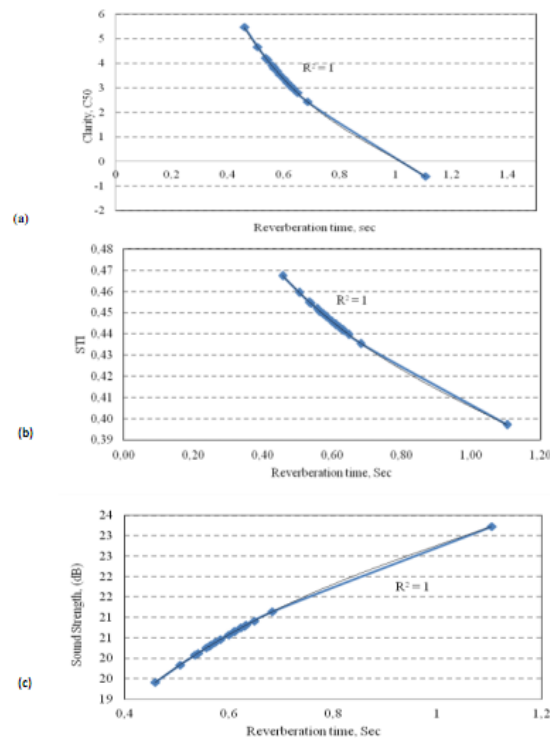
Неважко помітити, що графіки T20(f) і C50(f) тісно корельовані.

Значення індексу розбірливості STI наведені в таблиці:

Table 2 – STI values for classrooms

Classrooms	Volume, m ³	Average value of STI
P1	169	0.45
N1	220	0.46
Navet	575	0.50

Індекс STI був розрахований за спрощеною формулою Мюллера і Медіро [11]. Всі приміщення виявилися дуже посередніми за розбірливістю мови, оскільки за нормативами повинно бути $STI > 0,75$ (насправді, достатньо $STI > 0,6$ – А.П.). На представлених нижче графіках видно, що всі три показники C50, STI і G жорстко корельовані з часом реверберації T20 (коефіцієнт кореляції R практично дорівнює одиниці).



1.3.2.4. Di Marino C. et al. Acoustic Enhancement of Proposed Grand Lecture Hall using Computer Simulation [9]

У цій роботі наведені результати як комп'ютерного моделювання за допомогою програми ODEON, так і вимірювань акустичних параметрів, таких як час реверберації T60, EDT, STI і C80. Проблема полягала в тому, що досліджуваний великий лекційний зал (University of Windsor's Centre for Engineering Innovation (CEI) grand lecture hall) був уже побудований, проте остаточно не завершений. І хоча більш розумним було б подбати про його акустичні властивості ще на етапі проектування, ситуацію в якійсь мірі можна відкоригувати до завершення будівництва, використовуючи правильні покриття для стін і спеціальні акустичні фільтри.

Проблемою є те, що лекційний зал має розміри театрального залу, однак повинен володіти акустикою аудиторії. Для забезпечення надійності результатів, випробовувалися кілька приміщень до і після їх акустичного обладнання. Це дозволило правильно оцінити результати, отримані за допомогою ODEON. *В цілому результати досліджень ще раз показали, що акустичне проектування слід проводити на етапі загального проектування приміщень.*

Моделювання на ODEON дозволяло створити геометричні моделі приміщень. Таких моделей було дві: детальна і спрощена.

Результати моделювання та вимірювань показали, що програма ODEON давала результати з помітними помилками, що частково пояснюється відсутністю в її базі даних інформації про нові акустичні матеріали. Оцінка якості приміщення показала, що час реверберації є нижчим, ніж належить для приміщень таких розмірів. Пояснити це можна пірамідальної формою стелі і поглинанням поверхонь. В результаті лектор буде зазнавати труднощів під час відсутності системи звукопідсилення. Взагалі, спостерігається загальна тенденція робити приміщення більш заглушеними, що, однак, може створювати труднощі як для лекторів, так і для слухачів.

1.3.2.5. Bradley J., Reich R. and Norcross S. On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. - J.Acoust.Soc.Am., Vol. 106 (1999) [13]

У даній роботі проводилося оцінювання ряду показників, таких як відношення сигнал-шум, індекс STI і Alcons% при моделюванні різних сполучень шумових і ревербераційних умов. *Показано, що шумова перешкода набагато небезпечніша за реверберацію.*

Відзначено, що існує мало робіт, в яких одночасно досліджувалося б вплив шуму і реверберації. Причина в тому, що при дослідженнях легко варіювати значеннями сигнал-шум, але вельми важко варіювати рівнем реверберації. У даній роботі ця трудність подолана завдяки синтезуванню акустичних полів. Слухачі перебували в реальному тривимірному звуковому

полі, а управління SNR та реверберацією здійснювалося за допомогою комп'ютера.

При наявності тільки шуму, мірою розбірливості може служити артикуляційний індекс AI. При переважній дії реверберації мірою розбірливості може служити C50

$$C_{50} = 10 \log \left\{ \int_0^{0.05} p^2(t) dt / \int_{0.05}^{\infty} p^2(t) dt \right\}, \text{ dB}, \quad (1)$$

Що стосується мір, здатних одночасно враховувати дію шуму і реверберації, таких мір запропоновано кілька. Наприклад,

$$U_{50} = 10 \log \left\{ \frac{E}{L+N} \right\}, \text{ dB}. \quad (2)$$

де E – енергія ранніх відбиттів; L - енергія пізніх відбиттів; N – енергія шуму.

Зв'язок між U50 і C50 легко побачити, якщо переписати (2) у вигляді

$$U_{50} = 10 \log \left\{ \frac{E/L_{50}}{1 + (E/L_{50} + 1)N/S} \right\}, \text{ dB}, \quad (3)$$

та врахувати, що

$$C_{50} = 10 \log(E/L_{50}).$$

Інші міри, що одночасно враховують вплив шуму і реверберації, це Alcons% і STI. Оскільки математично ці три міри (U50, Alcons% і STI) описуються по-різному, неясно, наскільки однаково вони представляють дію шуму реверберації. Хоча недавні дослідження виявили сильну кореляцію між цими трьома мірами (JS Bradley, " Relationships among measures of speech intelligibility in rooms, " J. Aud. Eng. Soc. 46, 396-405 ~ 1998.), суб'єктивна оцінка розбірливості мови при цьому не проводилася. У даній роботі ця прогалина заповнюється. Крім того, випробовуються інші прості міри, які використовують комбінації SNR і часу реверберації або EDT. Взагалі кажучи, дослідження показують, що при одночасному врахуванні SNR і реверберації, оптимальні умови менш чутливі до дії реверберації, ніж це раніше вважалося.

Організація експериментів

Слухачі й тести. Використовувалися записи 4-х списків по 50 слів у кожному. Слухачі мали визначити тільки перші приголосні звуки кожного тестового слова. Тестові слова вимовлялися в несучій пропозиції типу «Слово номер ... це ... запишіть, будь ласка». Матеріал представлявся зі швидкістю 3 склади в секунду. У випробуваннях брали участь 10 осіб у віці від 20 до 62 років, які не мають проблем зі слухом. Кожен слухач тестувався індивідуально в створюваному акустичному полі, при цьому гучномовці та інше обладнання заглушеного приміщення не маскувалися. Кожен слухач прослуховував список з 50 слів в 16 різних акустичних умовах за один сеанс (поєднання 4-х значень SNR: -5, 0, +5 і + 10дБ і 4-х ревербераційних ситуацій - см табл. 2). Тести представлялися у випадковому порядку. Таких сеансів для кожного слухача було 4, всі сеанси проводилися в різні дні. В кінцевому рахунку оцінювався середній, по всім слухачам, відсоток правильно розпізнаних повідомлень.

Моделювання акустичних полів. Здійснювалося з використанням 8 гучномовців, розміщених на відстані 1,8 м від слухача в заглушеному приміщенні. Орієнтація гучномовців – в табл. 1. На гучномовці подавалися мовленнєві сигнали, навмисно спотворені реверберацією з необхідними властивостями, а також зашумлені з потрібним рівнем шуму. Формування сигналів вироблялося як за допомогою комп'ютера, так і за допомогою спеціалізованих цифрових пристроїв (еквалайзери, ревербератори, лінії затримки).

TABLE 1. Orientation of loudspeakers relative to the listener. Horizontally straight ahead of the listener at ear level is 0 deg in both planes.

Loudspeaker	Horizontal angle, deg	Vertical angle, deg
Center low	0	0
Center high	0	20
Left low	-32	0
Right low	+32	0
Left high	-37	23
Right high	+37	23
Far left	-90	0
Far right	+90	0

TABLE II. Measured mid-frequency (1 kHz) values of C_{50} , EDT, and RT for the four different room acoustics cases.

C_{50}	EDT	RT
-3.06	1.88	1.95
1.10	1.42	1.45
5.17	1.05	1.01
9.15	0.43	0.56

Результати експериментів

На Fig. 2 наведені результати оцінювання розбірливості для всіх 16 ситуацій. Добре видно, що фактор SNR набагато сильніше впливає на розбірливість мови, ніж фактор C_{50} . Дійсно при $\text{SNR} = 10$ дБ розбірливість мови вельми висока, і при зміні C_{50} в широких межах вона змінюється всього на 2,5%. Тим часом, коли показник C_{50} приймає високі значення, значення розбірливості мови при змінах SNR змінюються набагато більше, на 15% при $C_{50} = +9$ та на 25% при $C_{50} = +1$.

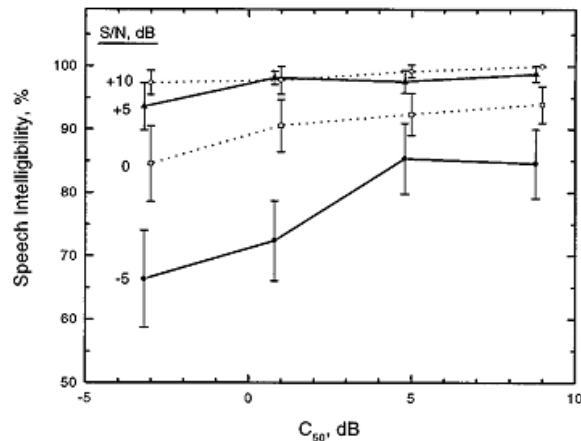


FIG. 2. Mean speech intelligibility scores for each of the 16 conditions and the corresponding 95% confidence limits. (The $\text{S/N} = +5$ dB and -5 dB cases have been offset to avoid overlapping error bars.)

В роботі також пропонується інший доказ порівняно більшого впливу шуму на розбірливість мови. З цією метою пропонується графік Fig.3, що зв'язує між собою об'єктивну і суб'єктивну міри розбірливості функціональною залежністю, добре апроксимується (коефіцієнт змішаної кореляції $R^2 = 0,973$) поліномом 3-го ступеня.

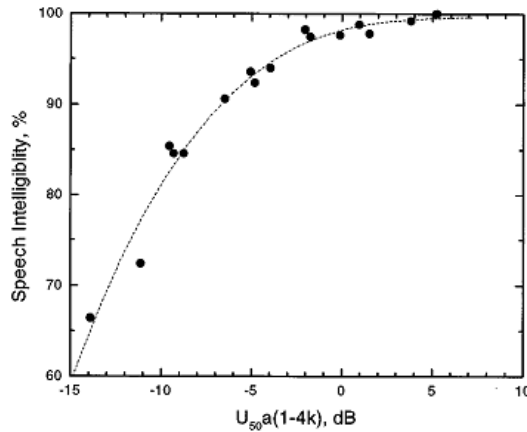


FIG. 3. Example of third order polynomial fit of broadband useful-to-detrimental ratio $[U_{50}a(1-4k)]$ to speech intelligibility scores.

Якщо окремо побудувати такі залежності для шумових і ревербераційних характеристик, виявляється, що для шумових об'єктивних показників $R2 = 0,78-0,86$, тоді як для ревербераційних об'єктивних показників $R2 = 0,14$.

Розробка нової широкосмужової міри відношення «корисне-шкідливе»

Результати обчислення показника $R2$ для різних варіантів нових мір наведені в табл. III. Нові міри базуються на співвідношеннях (2)–(3). Індекси при позначеннях мір показують, що перевірялися варіанти 50 та 80 мс кордону між корисною і шкідливою частиною ІХ приміщення. Індекс «а» означає, що інтенсивність шуму оцінювалася з А-зважуванням, а індекс в круглих дужках означає, що потужність оцінювалася у відповідній широкій смужі частот. Наприклад, індекс (1-4k) означає, що інтенсивність шуму оцінювалася за допомогою октавних фільтрів з центральними частотами від 1 до 4 кГц. До речі, саме така смуга забезпечує найкраще узгодження даних з поліноміальною кривою. Дана ситуація, що відповідає використанню співвідношення (2), зображена на графіку Fig. 3.

Ситуація, що відповідає використанню комбінованих мір (співвідношення (3)), представлено в табл. IV, де для зручності зіставлення наведені деякі дані з табл. III.

TABLE III. Results of third order polynomial fits of various broadband useful-to-detrimental ratio measures to speech intelligibility scores. In all cases there were 16 data points.

Measure	R^2	Measure	R^2
$U_{50a}(1-4k)$	0.973	$U_{80a}(1-4k)$	0.943
$U_{50a}(500-4k)$	0.969	$U_{80a}(500-4k)$	0.958
$U_{50a}(A)$	0.972	$U_{80a}(A)$	0.969
$U_{50a}(AI)$	0.971	$U_{80a}(AI)$	0.969
$U_{50}[C_{50}(500-4k)]$	0.961	$U_{80}[C_{80}(500-4k)]$	0.958

TABLE IV. Comparison of R^2 values from third order polynomial fits of combined measures to speech intelligibility scores.

Combined measure	R^2
$U_{50a}(A)$	0.972
$U_{50}[C_{50}(500-4k)]$	0.961
$U_{80a}(A)$	0.969
$U_{80}[C_{80}(500-4k)]$	0.958
STI+N	0.936
RASTI+N	0.738
AL _{cons} (2k)	0.930

Отримані результати зручно представити у формі, аналогічній Fig. 2. Наприклад, на Fig. 5 зіставлені результати експериментів (у вигляді точкових оцінок і довірчих інтервалів) з результатами прогнозування. На цьому рисунку добре видно, що при малих значеннях C_{50} і великих значеннях SNR прогнозні значення розбірливості виходять за довірчий інтервал, тобто не узгоджуються з результатами експериментальної суб'єктивної оцінки.

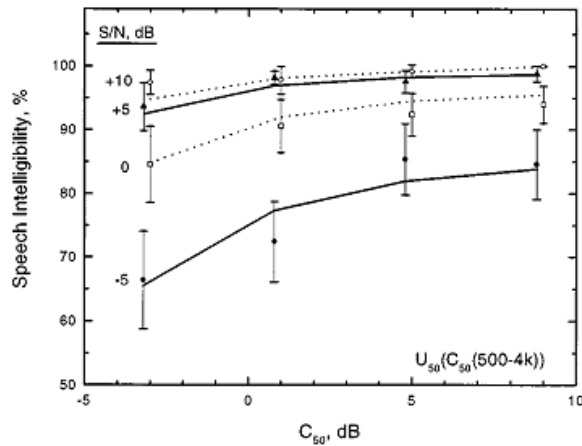


FIG. 5. Comparison of measured and predicted speech intelligibility scores based on the $U_{50}[C_{50}(500-4k)]$ measure. (The S/N = +5 dB and -5 dB cases have been offset to avoid overlapping error bars.)

1.3.2.6. Sato H., Bradley J. Evaluation of acoustical conditions for speech communication in working elementary school classrooms [14]

Представлені результати досліджень в 41 класному приміщенні Оттави. В результаті вимірювань виявлено, що середнє $SNR = 11$ дБА, $T60 = 0,3-0,6$ с. Крім того, досліджено вплив дітей на акустичну обстановку. У висновку написано, що *SNR є найбільш важливою характеристикою приміщень, особливо для дітей, що сидять далеко від викладача. Дитяча активність є значним додатковим джерелом шуму, підвищуючи середню SNR приблизно на 10 дБА.* Вельми сумнівно, щоб рівень шуму можна було знизити, використовуючи якісь звукопоглинальні покриття. Причиною підвищеного рівня шуму є прямий звук від ближніх дітей і знизити цей шум можна лише знижуючи рівень шуму. Як наслідок, рівень гучності викладача був вище, ніж очікувалося, і за шкалою Пірсона перевищував рівень «Підвищений». Очевидно, що викладач був змушений підвищувати гучність мови для того, щоб бути почутим на тлі навколишнього шуму. У розглянутих приміщеннях рівень гучності може бути підвищений на 5 дБ за рахунок збільшення частки ранніх віддзеркалень.

Деякі деталі досліджень

Досліджувалися класні кімнати 1, 3 і 6 класів. Середня кількість учнів у класі - 21. Сім приміщень були відкритого типу, решта 34 приміщення мали розміри 9.4m x 7.6m x 2.8m ($V = 197\text{m}^3$, $S = 236\text{m}^2$).

Результати вимірювань часу реверберації і C50 представлені в табл. 1. Як бачимо, в заповнених приміщеннях час реверберації на 10% менше, ніж в незаповнених. Відповідно, C50 на 1,34 дБА більше в заповнених приміщеннях.

Table 1: Mean reverberation times and early-to-late arriving energy ratios (C_{50}) measured in occupied and unoccupied classrooms.

Oct. band center frequency, Hz	125	250	500	1k	2k	4k	A-weighted
Occupied							
Mean rev. time,s	0.58	0.51	0.45	0.40	0.38	0.39	0.41
S.D.	0.14	0.09	0.10	0.11	0.09	0.08	0.09
Mean C_{50} , dB	5.34	6.39	7.98	9.75	11.12	11.46	10.49
S.D.	3.55	2.76	2.53	3.00	3.09	2.83	2.68
Unoccupied							
Mean rev. time,s	0.61	0.53	0.48	0.45	0.43	0.43	0.45
S.D.	0.15	0.10	0.11	0.12	0.12	0.11	0.11
Mean C_{50} , dB	5.20	6.01	7.37	8.32	9.58	9.87	9.13
S.D.	3.71	2.59	2.36	2.90	3.07	2.70	2.63

Середній рівень шуму в приміщеннях наведено в табл. 2.

Table 2: Background noise L_{eq} values measured in quiet occupied and unoccupied classrooms

Oct. band center frequency, Hz	125	250	500	1k	2k	4k	A-weighted
Occupied							
Mean	46.6	43.9	41.5	38.4	35.3	31.6	44.5
S.D.	6.2	4.0	3.9	4.1	3.7	3.3	3.7
Unoccupied							
Mean	46.0	42.1	39.6	36.2	31.6	26.7	42.1
S.D.	5.6	4.3	4.6	4.7	4.5	3.6	4.2

Як бачимо, рівень шуму перевищує рекомендований ANSI рівень 35 дБА. Рівні мови шуму вимірювалися на 200 мс інтервалах, після чого будувалися гістограми результатів вимірювань (Fig. 5).

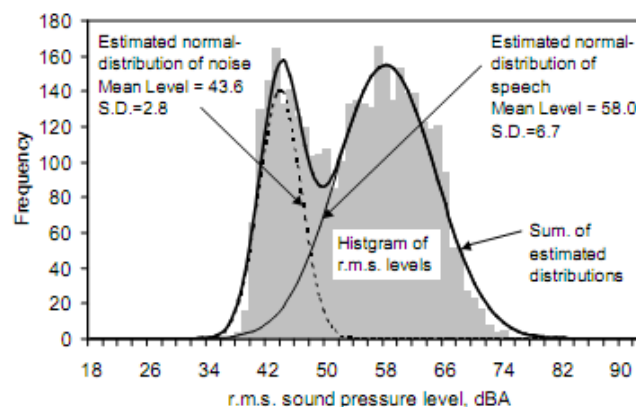


Figure 5: Example of histogram of r.m.s. sound pressure level with 200ms of time window obtained by an active class recording at one microphone position.

Гістограми результатів для 28 приміщень наведено на Fig. 6.

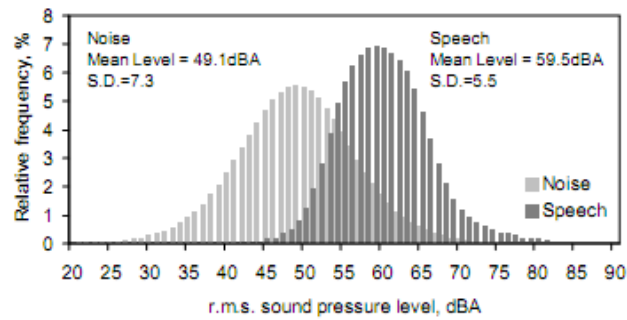


Figure 6: Relative frequency distribution of speech and noise level on the average of 28 classes recording.

Гістограма оцінок відношень сигнал-шум наведена на Fig. 7.

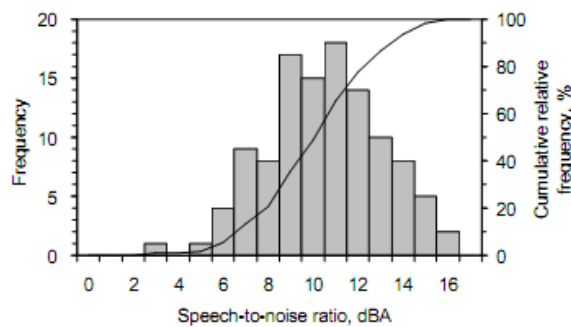


Figure 7: Frequency distribution of speech-to-noise ratio and its cumulative relative frequency sum up all of mean S/N measured in each measurement point ($N=118$). Mean S/N of all classrooms is 11dBA(S.D.=2.55) .

Що стосується гучності мови на відстані 1 м від диктора, вона була близькою до 69 дБА. За шкалою Пірсона, в 16% це відповідало позиції між градаціями «Нормальна» і «Підвищена», в 52% - між «Підвищена» і «Гучна», в 29% - між «Гучна» і «Крик».

1.3.2.7. J. S. Bradley, H. Sato: Speech intelligibility test results for grades 1, 3 and 6 children in real classrooms [15]

Це продовження представленої вище роботи [14]. Оцінювалася розбірливість мови для підлітків різного віку і для різних SNR. Використовувалися тести: WIPI (Word Intelligibility by Picture Identification),

Rhyme test, та Listening Difficulty test. Результати попередніх досліджень різних авторів представлені на рис. 1.

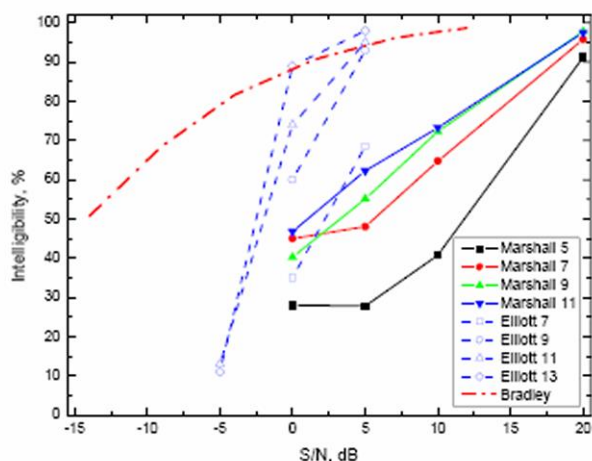


Рис. 1

Тест WIPI хороший тим, що зрозумілий навіть дитині 5 років: ведучий вимовляє фразу «Помітьте малюнок з ..., будь ласка». Використовувалися 4 мікрофона, біля кожного сиділи 5 підлітків. Так було протестовано 840 підлітків у віці 6, 8 і 11 років, в 41 приміщенні. Крім SNR, оцінювалися також характеристики приміщень: час реверберації, відношення енергій, відношення «корисний-шкідливий» та індекс STI.

Крім того, тестувалися молоді дорослі люди, щоб можна було впевнено говорити про вплив віку на розбірливість мови. Результати рис. 2 наочно демонструють залежність розбірливості мови від віку.

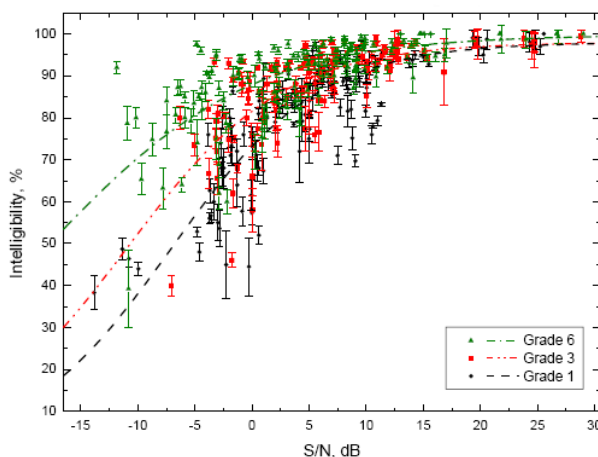


Рис. 2

Ці результати дещо по-іншому інтерпретовані на рис. 3, де добре видно, які значення SNR потрібні кожній віковій групі для досягнення розбірливості 95%. Добре видно, що різниця SNR для дітей 6 і 11 років склала 7 дБ. При дуже високих SNR (20-30 дБ) розбірливість досягає 98% для дітей 6 і 8 років, а для хлопців 11 років розбірливість досягає 99,5%.

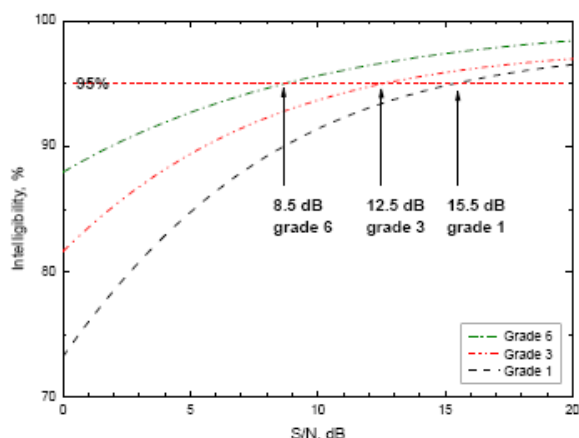


Рис. 3

На рис. 4 крім результатів дітей наведені результати молодих дорослих людей (суцільні помаранчева і синя лінії). Різниця між цими кривими і кривими дітей 11 років можна пояснити не тільки відмінністю віку, а й деяким розходженням умов тестування.

Проте, розбірливість 95% досягається дорослими вже при $SNR = +1$ дБ.

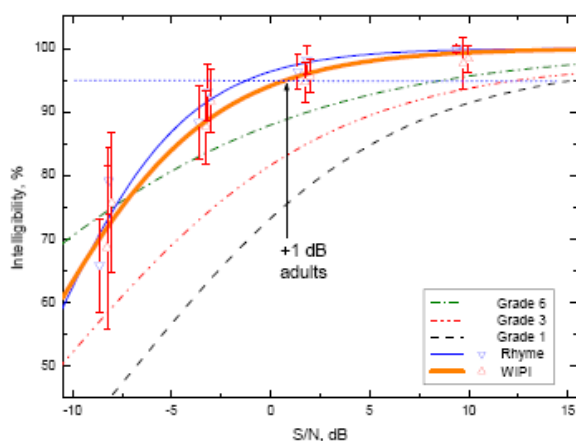


Рис. 4

Дорослі також тестувалися тестом «Труднощі сприйняття» - ці результати показані на рис. 5.

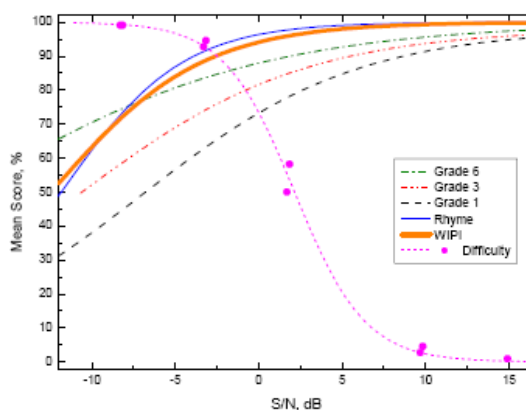


Рис. 5

1.3.2.8. Yang W. and Bradley J. Effects of room acoustics on the intelligibility of speech in classrooms for young children [16]

Це ґрунтовна (12 сторінок) і досить свіжа (2008-2009 рр.) стаття з JASA. Тому досить докладно розглянемо викладені в ній результати.

У Вступі згадуються результати попередніх робіт Bradley [14,15], де показано, що на розбірливість мови впливають три основні чинники: SNR, відбитий звук і вік слухачів. *У цих роботах досліджено 41 приміщення, однак час реверберації в них був практично одним і тим же, тому мета даної роботи полягає в заповненні цього недоліку, а саме - дослідження впливу часу реверберації на розбірливість мови при фіксованому SNR.*

У більшості робіт вважалося, що розбірливість зростає зі зменшенням часу реверберації. Однак є роботи (Bradley), в яких показано, що ранні відбиття можуть забезпечувати такий же ефект, як істотне збільшення рівня прямого звуку (7 дБ і більше). Hodgson and Nosal показали, що для ситуації, коли джерело шуму знаходиться ближче, ніж диктор, ранні відбиття можуть значно збільшити SNR.

Саме така ситуація має місце в класах, де основним джерелом шуму є учні, які сидять до слухача набагато ближче, ніж диктор. Позитивна роль ранніх віддзеркалень, що збільшують SNR, в даному випадку очевидна. Разом з тим,

слід пам'ятати про існування особливих ситуацій, коли ранні відбиття не сприяють підвищенню розбірливості. Це має місце, наприклад, коли диктор не повернутий особою до слухача або коли рівень прямого звуку менше рівня ранніх віддзеркалень (як це буває на великих дистанціях).

У даній роботі особливу увагу приділено вивченню залежності розбірливості мови від різних поєднань SNR і T60. При цьому розглянуті 2 типу сполучень SNR і T60. У першому випадку варіювалося T60, а SNR підтримувалося постійним. У другому випадку SNR збільшувалася за рахунок відображень при зростаючих значеннях T60. Нарешті, досліджувалося, як ранні відображення сприяють підвищенню розбірливості мови. Випробування проводилися на учнях у віці 6, 8 і 11 років, а також на дорослих слухачів.

Особливості експериментальних досліджень описані в частині 2 статті. У вимогах до акустики приміщень відзначено, що при моделюванні приміщень недостатньо варіювати часом реверберації, оскільки при цьому є небезпека отримати невірні результати через надто велику або занадто малу частку ранніх віддзеркалень. Наведено графіки залежностей між T60 і C50, що дозволяють врахувати дану обставину. Зазначено, що передбачається розглянути наступний набір значень T60: 0.3, 0.6, 0.9, і 1.2 с. Ці значення охоплюють весь діапазон значень T60, типових для шкільних класів і аудиторій.

Розглянуто дві групи ситуацій. У першій групі SNR підтримувалося постійним за рахунок того, що збільшення відбиттів однаковим чином позначалося на сигнальному і шумовому компонентах. У другій групі ситуацій рівень шуму підтримувався постійним, а рівень сигналу підвищувався через додавання відбиттів. В останньому випадку зростання рівня сигналу визначався відповідно до результатів досліджень Беранека, який встановив зв'язок між (EDT/V) і G_{mid} , де V - об'єм приміщення (м³), G_{mid} - відносний рівень сили звуку в приміщенні. Хоча Беранек не привів аналітичний вираз такої залежності, його можна отримати шляхом ручної апроксимації:

$$10 \log\{(EDT/V)10^6\} = G_{\text{mid}} + 16. \quad (1)$$

звідки отримуємо

$$G_{\text{mid}} = 10 \log\{(EDT/V)10^6\} - 16. \quad (2)$$

Для приміщень із середнім об'ємом 198 м³ краще підходив вираз

$$G_{\text{mid}} = 10.75 \log\{(EDT/V)10^6\} - 17.6. \quad (3)$$

котрий використовувався для оцінки приросту рівня сигналу при збільшенні EDT (див. Fig 2).

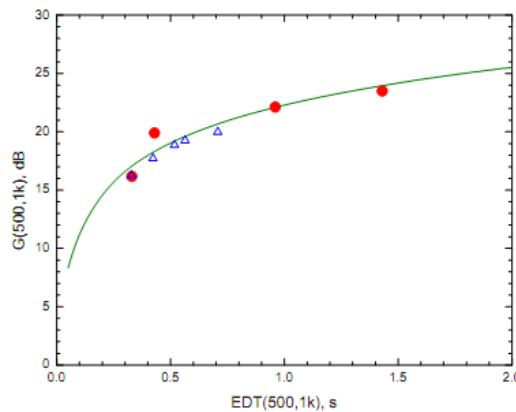


FIG. 2. (Color online) Variation in sound levels with decay time plotted as midfrequency G values vs measured midfrequency EDT values corresponding to the conditions with T_{60} values of 0.3, 0.6, 0.9, and 1.2 s for 198 m³ room. Solid circles: simulated conditions, open triangles: classroom-sized room data, and solid line: Eq. (3).

Слухачі і процедури випробувань. В якості мовного тесту використовувався WIPI. Слухачі сиділи в затишному оточенні, їм за допомогою портативного комп'ютера і навушників пред'являлися wav-файли (тестове слово, вимовлене всередині несучої фрази). Слухачі повинні були вказати курсором на потрібну (одну з шести) картинку, відповідну названому слову. Все це виглядало як комп'ютерна гра. В цілому тестувалися 234 людини (77 - 1-й клас, 75 - 3-й, 65 - 6-й і 17 дорослих).

Моделювання акустичних полів і застосування навушників. Випробування проводилися в заглушеному приміщенні. Звук випромінювався за допомогою 8 гучномовців Tannoy model 800A, що оточують слухачів. П'ять гучномовців знаходилися в одній горизонтальній площині, інші три розміщувалися фронтально трохи вище. Сигнали, що подаються на гучномовці,

оброблялися чотирма сигнальними процесорами Yamaha DME32, об'єднаними в один пристрій. Прямий звук формувався фронтальним гучномовцем. Ранні відбиття в кількості 31 формувалися і випромінювалися усіма 8 гучномовцями протягом 50 мс. Далі випромінювалася дифузна частина реверберації. *Час реверберації варіювався цифровим ревербератором на основі DME32.* Параметри C50 і T60 незалежно регулювалися в октавних смугах від 125 до 8000 Гц.

Для запису тестових сигналів використовувався акустичний манекен Brüel and Kjaer type 4128. Тестовий сигнал випромінювався 8 гучномовцями і записувався з виходів мікрофонів акустичного манекена. Потім ці записані сигнали прослуховувалися слухачами за допомогою навушників Sennheiser HD280 HP. *Оскільки спектр записаних сигналів спотворюється частотними характеристиками манекена і навушників, спектр випромінюваного звуку корегувався для компенсації цих спотворень.* Передатна характеристика ланцюжка «манекен–навушники» обчислювалася за результатами вимірювання ІХ навушників, одягнених на манекен.

Значна складність застосування навушників полягає в тому, що зміна положення навушників може призводити до істотної зміни передавальної характеристики навушників. Тому передавальна характеристика обчислювалася шляхом усереднення 10 передавальних характеристик, отриманих при різних положеннях навушників на голові манекена. При цьому спочатку усереднювалися ІХ, а потім вже обчислювалася середня передавальна характеристика.

Записаний мовний матеріал коригувався шляхом деконволюції записів з усередненою ІХ, щоб усунути вплив ланцюжка «манекен–навушники».

Навколишній шум був записаний аналогічним чином. Спектр шуму спадав зі швидкістю 5 дБ/октаву, що є типовим для шуму всередині приміщень, обумовленого дією вентиляції. Шум був бінауральним. Записи шуму змішувалися з записами мови для встановлення необхідного SNR.

Умови, в яких проводилися випробування. Розглянуто 9 умов випробувань (табл. II).

TABLE II. Description of the nine acoustical conditions used in the speech tests.

Condition	T_{60} (s)	Speech level (dBA)	Noise level (dBA)	S/N(A)
1	0.3	62	67	-5
2	0.6	65	67	-2
3	0.9	67	67	0
4	1.2	69	67	2
5	Direct only	60	67	-7
6	Direct+early	66	67	-1
7	0.3	62	64	-2
8	0.9	67	69	-2
9	1.2	69	71	-2

Експеримент 1 проводився в умовах 1-4. Експеримент 2 проводився в умовах 7, 2, 8 і 9. Експеримент 3 проводився в умовах 5, 6 і 3.

Оцінка якості навушників. Мета даної оцінки - з'ясувати, що за допомогою навушників можна досить точно моделювати прямий звук в заглушеному приміщенні. Дослідження показали, що таке моделювання має право на існування.

Результати експериментів описані в частині 3. В 1-м експерименті варіювалися значення SNR за рахунок варіювання T_{60} . Результати показані на Fig. 6.

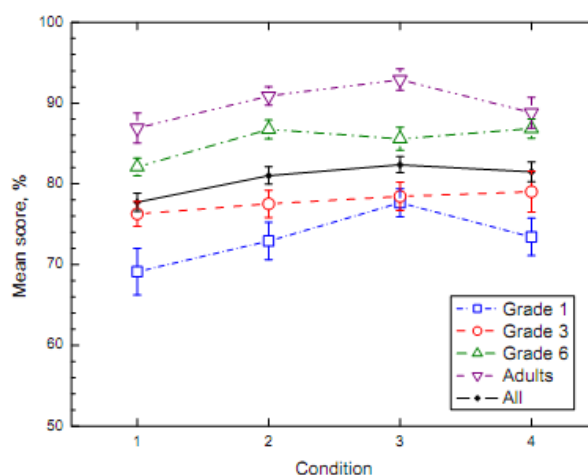


FIG. 6. (Color online) Mean corrected scores for conditions 1-4 having T_{60} values of 0.3, 0.6, 0.9, and 1.2 s, respectively. Each line refers to the data from a different age group and the error bars are the standard errors of each mean value. A fifth line indicates the averages of all four age groups.

Тут SNR збільшувалася не стільки за рахунок збільшення рівня мовного сигналу, скільки за рахунок збільшення часу реверберації. Добре видно вплив віку слухачів на розбірливість мови.

У 2-му експерименті значення SNR були постійними, хоча T_{60} варіювалося в тих же межах, від 0.3 до 1.2 с. Результати показані на Fig. 7.

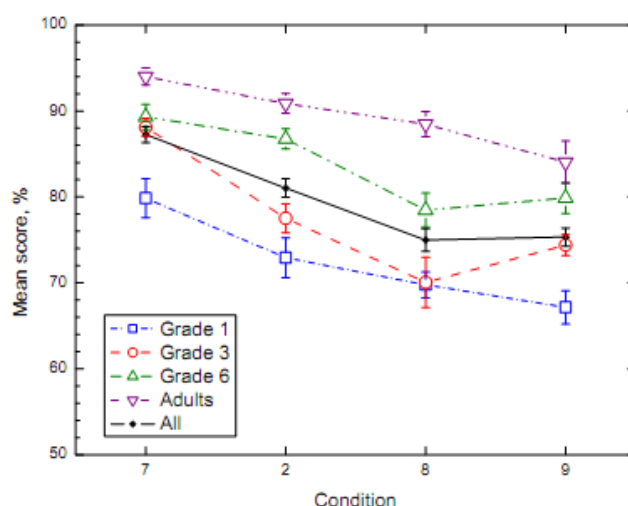


FIG. 7. (Color online) Mean corrected scores for conditions 7, 2, 8, and 9 having T_{60} values of 0.3, 0.6, 0.9, and 1.2 s, respectively, and a constant $S/N(A) = -2$ dB. Each line refers to the data from a different age group and the error bars are the standard errors of each mean value. A fifth line indicates the averages of all four age groups.

Бачимо, що при сталості SNR зростання T_{60} призводить до погіршення розбірливості, оскільки відсутній вигравш за рахунок відбить.

У 3-му експерименті перевірялася роль ранніх і пізніх відбить. Результати показані на Fig. 8. Видно, що наявність ранніх відбить істотно підвищує розбірливість, тоді як пізні відбиття негативно позначаються на розбірливості мови.

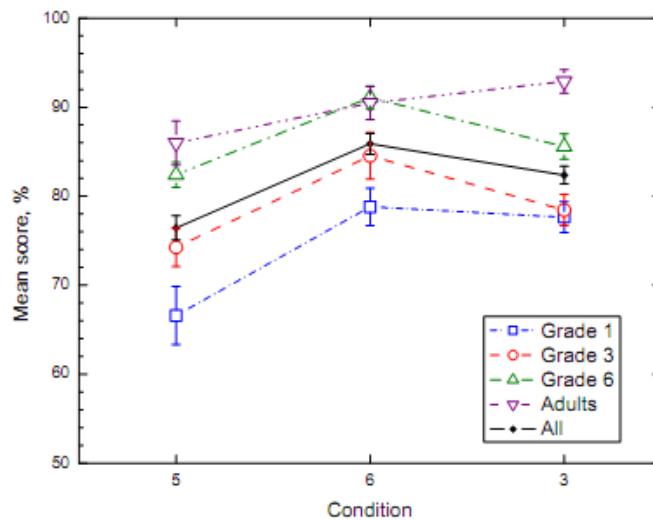


FIG. 8. (Color online) Mean corrected scores for condition 5 (direct sound only), condition 6 (direct sound and early-arriving reflections), and condition 3 (direct sound with early and late-arriving reflections). Each line refers to the data from a different age group and the error bars are the standard errors of the mean values. A fifth line indicates the averages of all four age groups.

Обсуждение полученных результатов.

Автори намагаються визначити ідеальні умови для комунікації в класах. З цією метою на графіку Fig. 11 представлені залежності розбірливості від U_{50} , на основі яких потім побудовані (Fig. 12) залежності розбірливості від T_{60} . Ці графіки відповідають результатам експерименту №1, коли SNR змінювалися через варіювання T_{60} . Неважко бачити (Fig. 12), що *оптимальним є значення $T_{60} = 0,68$ с, хоча екстремум виражений слабо, що свідчить про слабкий вплив реверберації на розбірливість. Значно яскравіше виражена залежність розбірливості від віку, причому кривизна лінії для наймолодших школярів найбільша, що свідчить про більшу чутливості їх слуху до впливу часу реверберації.*

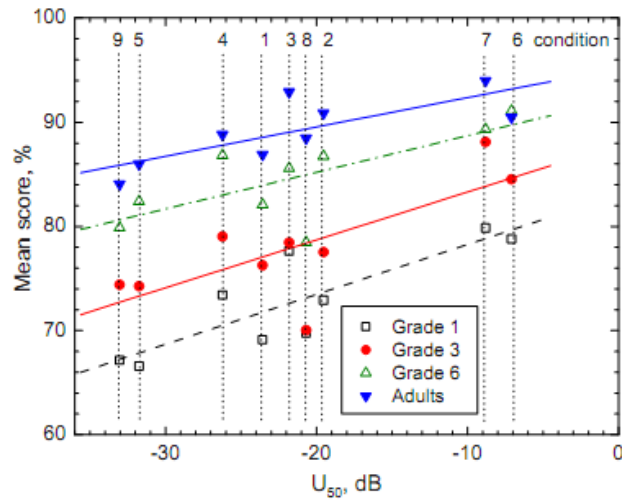


FIG. 11. (Color online) Plot of mean speech intelligibility scores vs U_{50} values for each of the nine conditions and for each age group with associated linear regression lines for each age group. Each vertical dotted line indicates the data for one condition as labeled at the top of the graph.

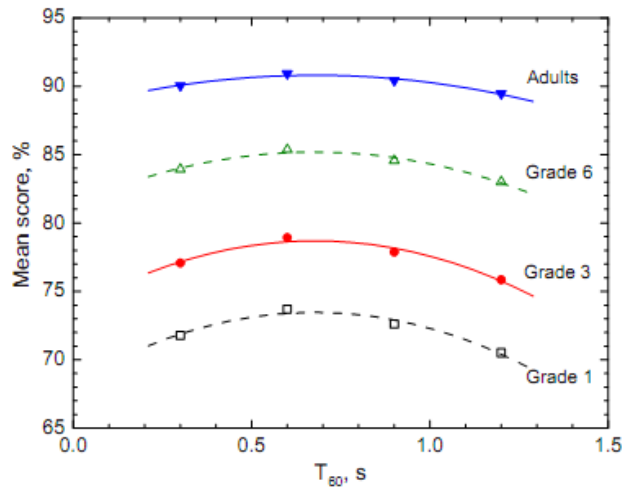


FIG. 12. (Color online) Smoothed speech intelligibility scores plotted vs T_{60} values for the results of experiment No. 1 with conditions having varied S/N values. The curved lines are second order polynomial regression lines to the data.

Експерименту №2 відповідає Fig. 13, з якого випливає, що при підтримці постійного SNR і варіюванні T_{60} оцінки розбірливості мови тим вище, чим менше час реверберації. Тут також добре видно, що зростання часу реверберації найбільш негативно позначається на школярах молодших класів.

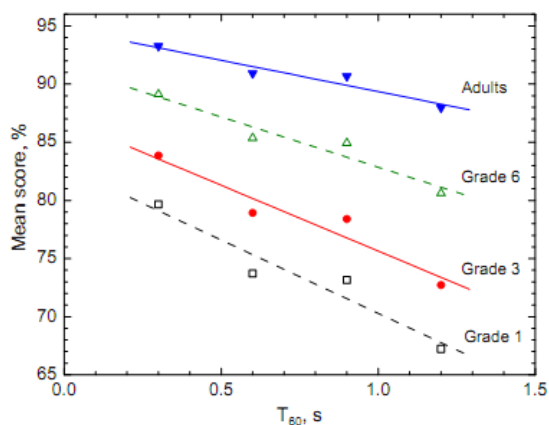


FIG. 13. (Color online) Smoothed speech intelligibility scores plotted vs T_{80} values for the results of experiment No. 2 with conditions having constant S/N value. The lines are linear regression lines to the data for each age group.

У висновках по роботі відзначається, що при проектуванні класних приміщень слід забезпечувати: 1) мінімальну гучність; 2) оптимізацію часу реверберації (0.3-0.9 с) для використання корисного вкладу відбить.

1.3.2.9. Bradley J.S. Review of objective room acoustics measures and future needs [17]

У цій роботі аналізуються об'єктивні міри якості приміщень, призначених для слухання музики. Розглянуто проблеми, пов'язані із застосуванням існуючих мір. Крім того, розглянуто ряд нових, недавно запропонованих мір і сформульовані нові завдання.

1.3.2.10. Висновки за оглядом

Принципова відмінність акустичної паспортизації приміщень, що проектуються [1-6], від такої для готових приміщень [7-18] полягає в тому, що вона спирається на результати розрахунків і моделювання. Продуктом акустичної паспортизації готових приміщень є результати вимірювань акустичних характеристик цих приміщень.

В обох випадках ключовим є перелік характеристик, що визначають кінцеві, тобто «споживчі» властивості приміщення. Тому немає нічого дивного в тому, що такі переліки визначаються функціональним призначенням приміщень.

Так, наприклад, загальна кількість вимірюваних за допомогою програми EASERA [19] характеристик і параметрів приміщення перевищує два десятки. Ці параметри згруповані у вигляді таких переліків:

- параметри загального призначення;
- параметри, пов'язані з позиціонуванням слухачів при мовних презентаціях;
- параметри, пов'язані з позиціонуванням слухачів при виконанні музики;
- параметри, що характеризують сприйняття якості приміщення слухачами;
- параметри, що характеризують сприйняття якості приміщення музикантам і диригентом.

Численні і розрізнені відомості про статистичну та функціональну залежності оцінок різних акустичних параметрів зібрано в роботі Д. Бредлі [18], де показано, що все різноманіття акустичних характеристик приміщень, призначених для музичних презентацій, можна звести до чотирьох-п'яти інформативних параметрів. Через 20 років, в роботі [17] цим же автором проаналізовано проблеми, пов'язані із застосуванням існуючих об'єктивних мір якості приміщень, призначених для слухання музики, а також зроблена спроба прогнозу потенційних можливостей ряду нових, недавно запропонованих мір.

Література до розділу 1.3

1. «Акустический паспорт» залов и помещений.
<http://www.smspro.ru/artic/acoustic-passport.php>
2. Акустическая экспертиза.
http://www.aspd.ru/napravlenija_dejatelnosti/akusticheskaya-expertiza/
3. Акустическое проектирование.
<http://www.tecsound.com.ua/proektuvannya/akustichne-proektuvannya/>
4. Акустический проект. <http://www.airfon-media.ru/acoustical-project/>
5. Акустика помещений.
<http://www.acoustic.ua/recommendations.html?g=3>

6. Алдошина И., Приттс Р. Музыкальная акустика. – Композитор, С-Пб, 2006. – 719 с.
7. Eggenschwiler K. Lecture Halls - Room Acoustics and Sound Reinforcement. - Forum Acusticum 2005 Budapest. - P. 1-6.
8. Youssef R., Bard D., Mahmoud F., Esa N. Acoustical Quality Assessment of Lecture halls at Lund University, Sweden. - Inter-noise, Melburn, Australia, 16-19 November, 2014. - P. 934-943.
9. Di Marino C. et al. Acoustic Enhancement of Proposed Grand Lecture Hall using Computer Simulation. - Canadian Acoustics, Vol. 39, No. 1, 2011. - P. 43-48.
10. ISO – 3382-1, Acoustics – “Measurements of room acoustic parameters” – Part 1: Performance spaces, 2009.
11. Muller, Swen Mediro STI (2005).
12. Усик В., Зубченко Д. Разработка аппаратно-программного комплекса для измерения объективных параметров акустики помещения. - Технология приборостроения, специальный выпуск, 2014. - С. 36-40.
13. Bradley J., Reich R. and Norcross S. On the combined effects of Signal-to-noise ratio and room acoustics on speech intelligibility. - J.Acoust.Soc.Am., Vol. 106 (1999).
14. Sato H., Bradley J. Evaluation of acoustical conditions for speech communication in working elementary school classrooms. - J. Acoust. Soc. Am. 106 (4), Pt. 1, May 2004. - P.2064-2077.
15. J. S. Bradley, H. Sato: Speech intelligibility test results for grades 1, 3 and 6 children in real classrooms, Proceedings of ICA, Kyoto (2004).
16. Yang W. and Bradley J. Effects of room acoustics on the intelligibility of speech in classrooms for young children. - J. Acoust. Soc. Am. 125 (2), February 2009. - P. 922-933.
17. J.S. Bradley Review of objective room acoustics measures and future needs. - Proceedings of the International Symposium on Room Acoustics, ISRA 2010, Melbourne, Australia, August 29-30, 2010, pp. 1-8.

18. J.S. Bradley. The Evolution of Newer Auditorium Acoustics Measures. - Canadian Acoustics, Vol.18, No.4, pp. 13-23. – 1990.
19. Ahnert W., Schmidt W. Appendix to EASERA Manual: Fundamentals to Perform Acoustical Measurements. – Berlin, 2006. – P. 1 – 53.
20. Васильев А.В., Климков В.А., Продеус А.Н. Автоматизация акустической экспертизы помещений. - Electronics and Communications 2013, №5(76). - С. 63-71.

Лекція 1.4. Джерела тестових сигналів: переваги та недоліки

Нижче наведено огляд літературних джерел, де розглянуто питання порівняння джерел тестових сигналів, призначених для вимірювання ІХ приміщень.

1.4.1. M.J.R. Lamothe and J.S. Bradley, "Acoustical characteristics of guns as impulse sources," Can. Acoust. 13(2), 16-24 (1985).

Виміряно акустичні характеристики п'яти револьверів трьох калібрів та двох типів для визначення їх придатності як джерел імпульсу. Порівняння проводили із джерелами іскрових та гучномовних імпульсів. Було виявлено, що револьвер калібру 0,38 є придатним повторюваним всенаправленим імпульсним джерелом, що забезпечує адекватну звукову енергію в октавних діапазонах від 125 до 5000 Гц.

Типові вимоги до джерел. В якості джерел звуку використовують револьвери, іскрові прилади, резинові кулі та гучномовці.

Механічні: Джерело має бути легким та портативним.

Акустичні: Всеспрямованість, рівність АЧХ, забезпечує велике SNR та повторюваність, короткість імпульсу. ІХ зазвичай вимірюють для октавних смуг від 125 Гц до 8 кГц, інтервал спадання має бути близьким до 30-40 дБ.

Увага! Не дуже хорошу рівність АЧХ можна компенсувати за умови, що така нерівність є гладкою.

Увага! Не дуже хорошу повторюваність можна компенсувати усередненням результатів, проте треба пам'ятати, що Т60 та С50 є чутливими до нестабільності джерела.

Увага! Проблема відсутності всеспрямованості утруднює порівняння отриманих результатів з іншими результатами.

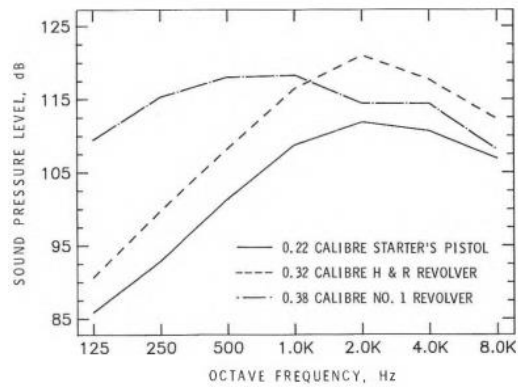


Figure 2. Octave band spectra of guns at 0 degrees incidence

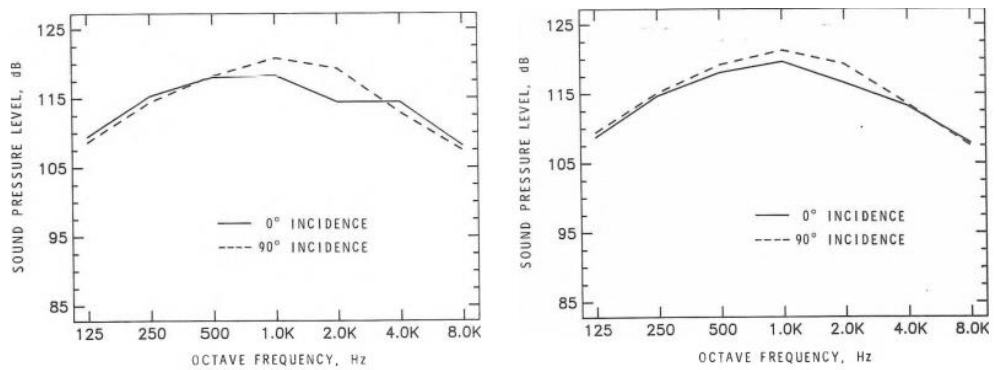


Figure 3-4. Octave band spectra of 0.38 calibre No. 1 revolver in horizontal (a) and vertical (b) plane

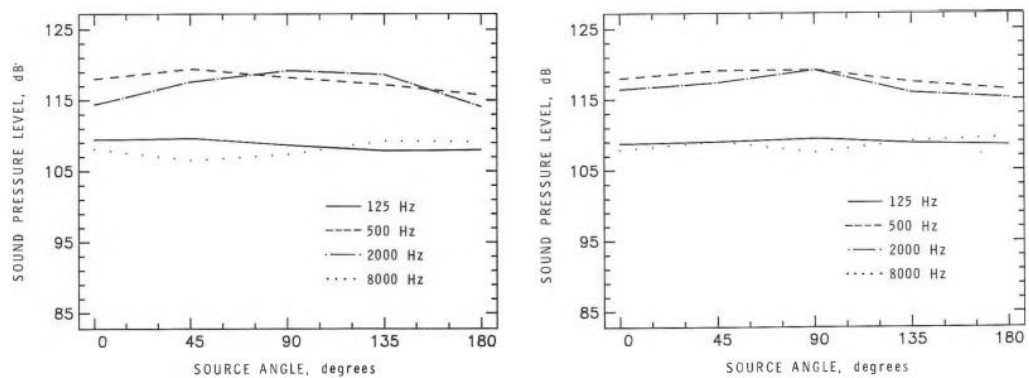


Figure 7-8. Octave band spectra levels in the versus angle of 0.38 calibre No. 1 revolver in horizontal (a) and vertical (b) plane

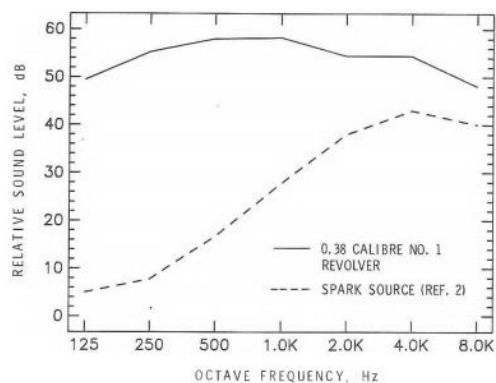


Figure 10. Comparison of octave band gun and spark source.

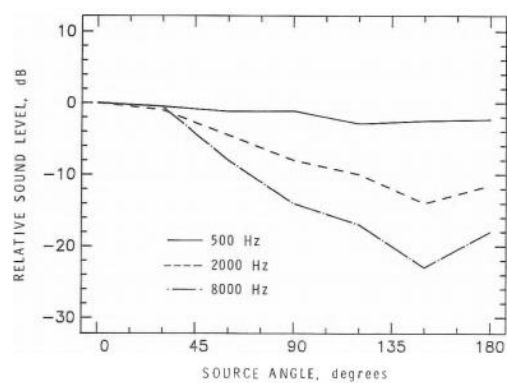


Figure 11. Octave band sound levels versus angle for a small loudspeaker.

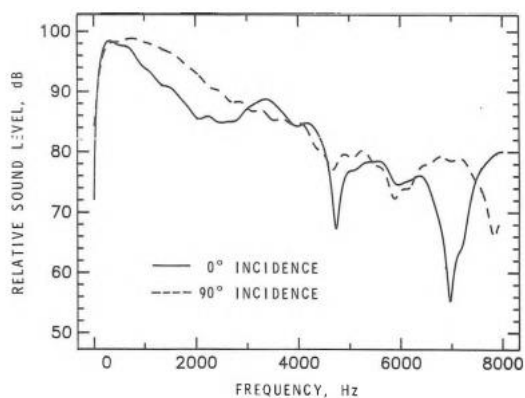


Figure 12. Narrow band power spectrum of 0.38 calibre No. 1 revolver.

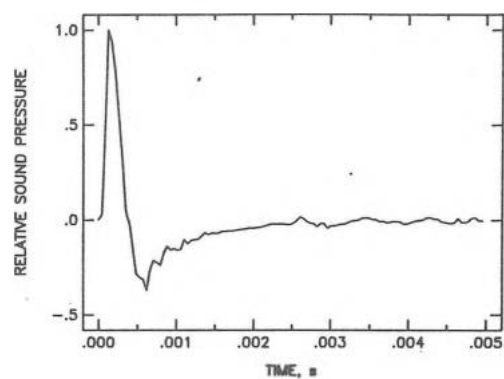


Figure 13. Anechoic time history of 0.38 calibre gun.

1.4.2. Bradley, J. S. Auditorium acoustics measures from pistol shots. *Journal of the Acoustical Society of America*, 80, 1, pp. 199-205, 1986-07

В статті розглянуто ряд монофонічних кількісних характеристик приміщення при записі звуку пострілу. Представлено результати вимірювань в 5 приміщеннях. Характеристики, що вимірювалися: RT, EDT, C80, TS, STI. Всі оцінки, за винятком RT, помітно змінюються із зміною місцеположення джерела та приймача звуку. Для обробки великої кількості даних важливо мати комп'ютеризовану систему. Вимірювальна апаратура має бути компактною та має дозволяти вимірювання швидко.

Бажано, щоб спектр короткого імпульсу був рівномірним, а джерело звуку було всеспрямованим на всіх частотах (в октавних смугах від 125 Гц до 8 кГц). Експеримент показали, що підходить пістолет калібру 0.38 (Fig. 1).

Обґрунтування такого вибору наведено в роботі M.J.R. Lamothe and J.S. Bradley, "Acoustical characteristics of guns as impulse sources," Can. Acoust. 13(2), 16-24 (1985). Гучномовці є гіршими в цьому сенсі, не дивлячись на існування різних способів боротьби із залежністю частотної характеристики від кута. Тому задовільні результати із гучномовцями можна отримати лише для вузьких частотних ділянок.

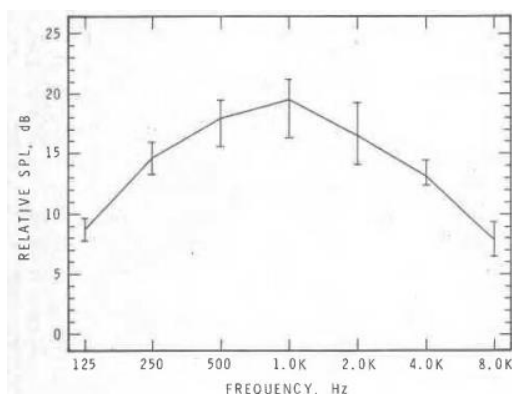


FIG. 1. Octave band spectrum of 0.38 caliber gun in an anechoic room for 0° incidence and range of values for other angles.

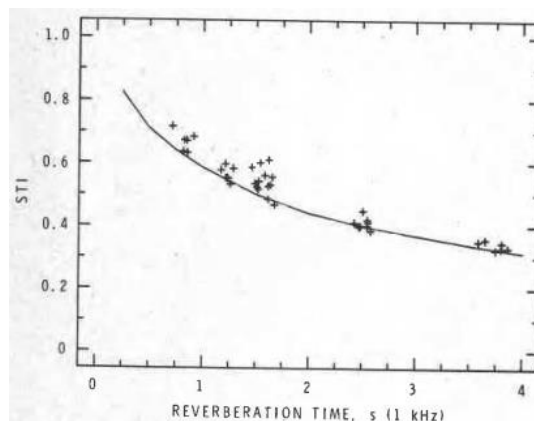


FIG. 6. Measured + + +; calculated —; STI vs RT.

Наявність магнітофону Technics SVP-10 із динамічним діапазоном 86 дБ та пласкою частотною характеристикою від 2 кГц до 20 кГц зробила записи звуку якісними.

Результати оцінювання STI в залежності від часу реверберації RT показано на Fig. 6.

Аналіз показав, що всі розглянуті показники є залежними між собою.

Стаття завершується побажанням порівняння розглянутих об'єктивних показників із результатами суб'єктивного оцінювання.

1.4.3. Bradley, J. S. Optimizing the decay range in room acoustics measurements using maximum-length-sequence techniques. Journal of the Audio Engineering Society, 44, (4), pp. 266-273, April 01, 1996

В статті розглянуто питання оптимізації вимірювань із застосування mls-послідовності.

Перевагою mls-послідовності є стійкість до фонового шуму. Проте для отримання значного динамічного діапазону відображення спадання звуку

необхідно оптимізувати вимірювальну систему. Для вимірювань потрібно максимізувати *інтервал спадання (decay range)* звуку – інтервал між максимумом I_X та її значенням в момент, коли спадання зрівнюється із фоновим шумом. Для такої максимізації треба або використовувати довгу mls-послідовність та/або робити усереднення результатів для кількох mls-послідовностей.

Обладнання, що використовувалося. Софт RAMSoft-3. Mls-сигнал генерувався програмно та випромінювався після 16-бітового ЦАПу через всеспрямованій гучномовець-додекаедр. Сигнал сприймався 2-ма мікрофонами – конденсаторним всеспрямованим та студійним із 8-діаграмою спрямованості для оцінювання латеральних відбиттів. Для створення фонового шуму контрольованого рівня використано електроакустичне обладнання для випромінювання синтезованого шуму поблизу мікрофонів (Fig.1). В статті говориться про генерування рожевого шуму, мотивуючи тим, що шум тестового сигналу є також рожевим. При цьому згадуються 1/3-октавні фільтри... Місцеположення джерела звуку та мікрофонами не змінювалося. Час реверберації регулювався шляхом зняття поглинального покриття зі стін кімнати. Довжина mls-послідовності була 16к, 32к, 64к, без одиниці (точніше, $2^{14}=16384$, $2^{15}=32768$ та $2^{16}=65536$), що для частоти дискретизації 12500 Гц становило 1,31, 2,62 та 5,24 с. Кількість усереднень результатів варіювалася від 1 до 100.

Використовувалося 4 різних види еквалізації тестових сигналів. Необхідність еквалізації спричинена нерівномірністю АЧХ гучномовця, а також необхідністю випромінювати, як варіант, рожевий шум. Варіант рожевого шуму пояснюється тим, що цим досягається випромінювання більшої потужності на НЧ, оскільки зазвичай шум оточення є більш потужним на НЧ, а також тим, що спектральний аналіз може виконуватися гребінкою 1/3-октавних або октавних фільтрів. При цьому відмічається, що підйом рівня сигналу на тих частотах, де АЧХ гучномовця є низькою, призводить до спотворень сигналу на виході

гучномовця (!!!). Це призводить до підвищення рівня шуму та зменшення інтервалу спадання.

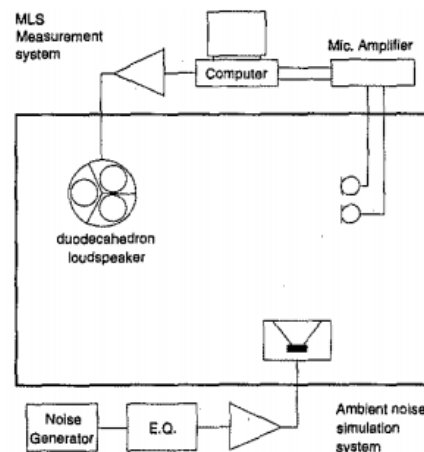


Fig. 1. Block diagram of test equipment showing MLS measurement system and ambient noise simulation system.

Графіки C50 для різних видів еквалізації наведено на Fig.2.

Погано також знижувати рівень сигналу на тих частотах, де він підсилюється гучномовцем – при цьому знову таки зменшується рівень сигналу і як наслідок, зменшується інтервал спадання. Тому в даній роботі використовувався апаратний еквалайзер (входить до складу гучномовця), застосування якого давало рожевий шум випромінення та лише невелике зниження інтервалу спадання.

На Fig.3 показано, що збільшення рівня сигналу призводить до збільшення інтервалу спадання, доки не починаються спотворення сигналу. На Fig.4 та 5 показано залежність інтервалу спадання від кількості усереднень та довжини mls-послідовності.

А на Fig.6 та 7 показано залежність інтервалу спадання одночасно від кількості усереднень та довжини mls-послідовності. Ці рисунки відрізняються рівнем вихідного сигналу. Цікаво, що в найкращому випадку вдається досягти лише 40 дБ інтервалу спадання.

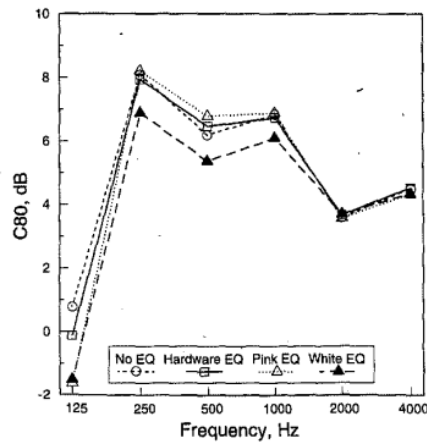


Fig. 2. Measured early-to-late sound ratios (C80) for four different source spectra: — no equalization; — hardware equalization; ··· pink equalization; — white equalization.

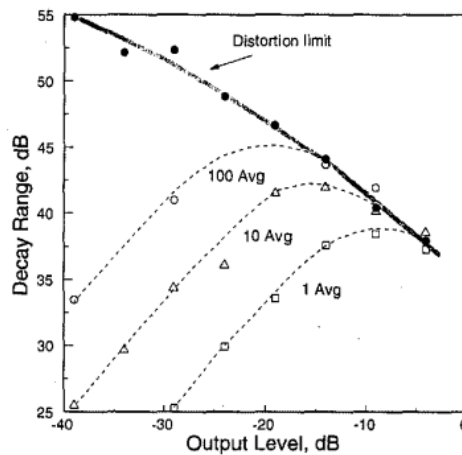


Fig. 4. Measured decay range versus system output level (relative to maximum system output) for various numbers of averages. All results were responses to 64k-point sequences over octaves from 500 to 4000 Hz and with 71-dB synthesized ambient noise.

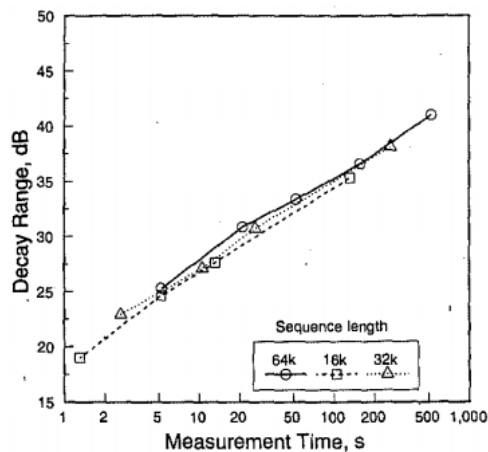


Fig. 6. Measured decay range versus total measurement duration for 16k-, 32k-, and 64k-point sequence lengths. System output level -29 dB. All measurements were averages over octaves from 500 to 4000 Hz with 71-dB synthesized ambient noise.

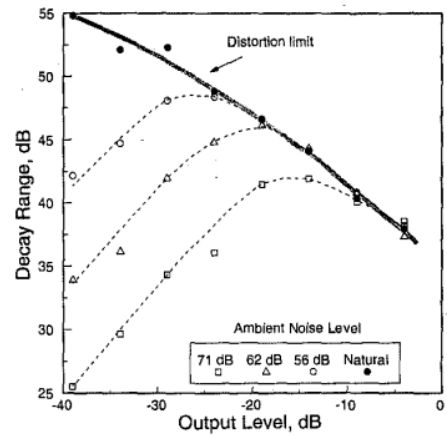


Fig. 3. Measured decay range versus system output level (relative to maximum system output) for various background noise levels. All results were obtained using 64k-point sequences and were averages of 10 sequences over octave bands from 500 to 4000 Hz.

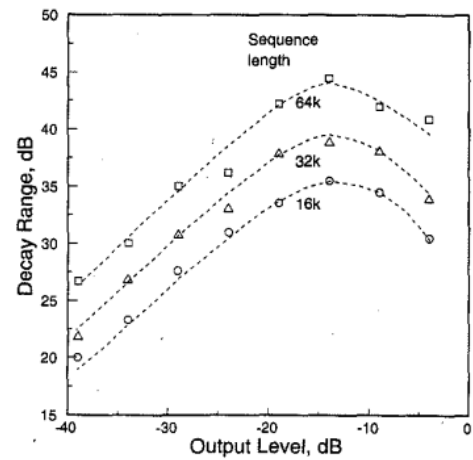


Fig. 5. Measured decay range versus system output level (relative to maximum system output) for 16k-, 32k-, and 64k-point sequence lengths. All measurements were averages of responses to 10 sequences over octaves from 500 to 4000 Hz and with 71-dB synthesized ambient noise.

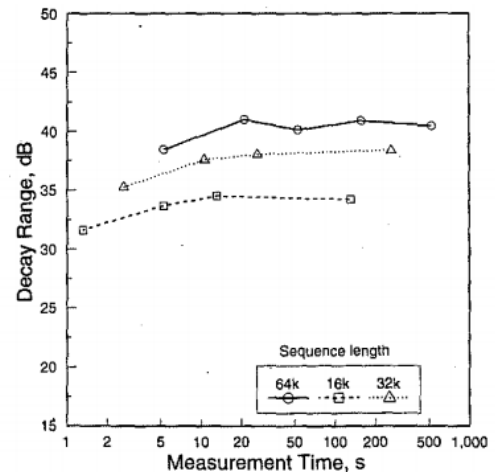


Fig. 7. Measured decay range versus total measurement duration for 16k-, 32k-, and 64k-point sequence lengths. System output level -9 dB. All measurements were averages over octaves from 500 to 4000 Hz with 71-dB synthesized.

Крім кількості усереднень та довжини mls-послідовності, є й інші фактори, що впливають на результати вимірювань. В даній роботі вказано на 2 таких фактори: зміна температури повітря в кімнаті та зміна часу реверберації. На рис. 8 показано, що час реверберації стає помітно меншим в період дії нагрівача, але після виключення нагрівача знову вертається до попередніх значень. Дія нагрівача також відбивається на параметрі C80 (рис. 9). Вплив зміни часу реверберації показано на рис. 10.

Практичні рекомендації:

- 1) визначте умови появи апаратних спотворень при низькому рівні шуму оточення;
- 2) при вимірюваннях в приміщенні, спершу визначте рівень шуму оточення;
- 3) використовуйте помірний рівень вихідного сигналу;
- 4) використовуйте mls-послідовність максимально можливої довжини;*
- 5) використовуйте усереднення результатів, якщо відсутні спотворення в гучномовці;
- 6) скорочуйте час вимірювань, якщо властивості приміщення змінюються із плином часу.

*Сучасні комп'ютери та програми дозволяють генерувати mls-послідовності довжиною до 2^{18} та вище.

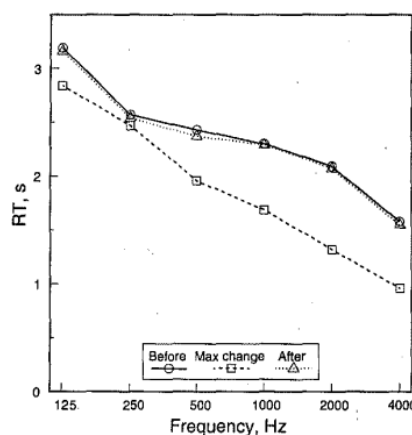


Fig. 8. Measured RT versus frequency before and after operation of a 1-kW heater as well as maximum change.

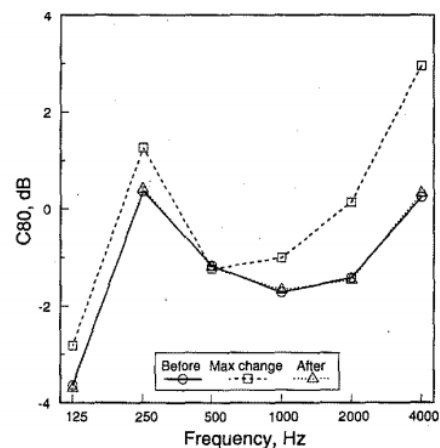


Fig. 9. Measured C80 versus frequency before and after operation of a 1-kW heater as well as maximum change.

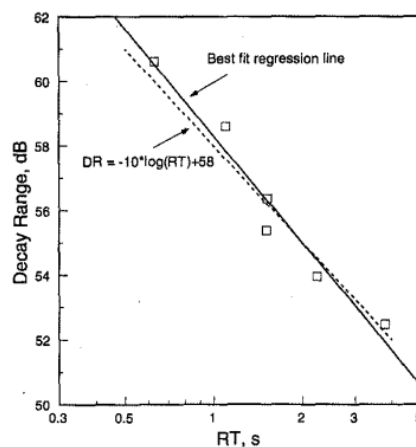


Fig. 10. Measured decay range DR versus RT measured on digital reverberator. All measurements were made using a 64k-point sequence.

ПРИМІТКА: В Матлабі (версія 2018b, не нижче) дуже просто генерувати mls-послідовність потрібної довжини, рівня та кількості повторювань: <https://www.mathworks.com/help/audio/ref/mls.html>

Рис. 11 допоможе в налаштуванні вимірювальної системи. Слід враховувати, що цей рисунок відповідає ситуації, коли довжина mls-послідовності становить 2^{16} , а кількість усереднень дорівнює 10. Тут «відносний рівень шуму оточення» означає наступне. Якщо відносний вихідний рівень сигналу становить 0 дБ для рівня звуку на вході мікрофона 100 дБ, тоді «відносний рівень шуму оточення» -40 дБ відповідає 60 дБ шуму оточення.

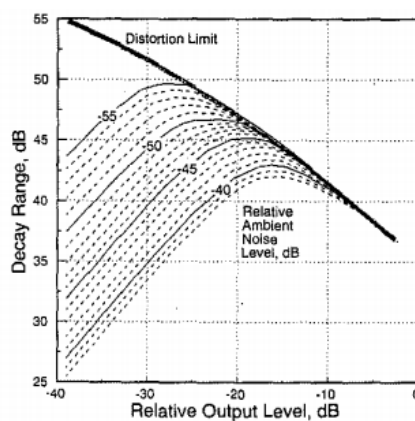


Fig. 11. System optimization chart in terms of decay range versus relative output level for various ambient noise levels. Output level and ambient noise level are both relative to level of output signal corresponding to maximum system output.

1.4.4. D. Griesinger, “Beyond MLS - Occupied hall measurement with FFT techniques,” in AES 101th Convention, Preprint 4403, Los Angeles, CA, Nov. 1996.

Стаття містить довгу низку програм для Matlab. Матеріалу багато (23 с.), є посилання на роботу Бредлі 1996 року (див. вище).

Дана стаття є цікавою із наступних причин:

- 1) оригінальна проблематика – оцінювання якості концертним залів для музики за наявності оркестру та публіки;
- 2) виконано власне порівняння просторової спрямованості та спектрального складу різних джерел звуку: холостий постіл із пістолету, гумова кулька, хлопок в долоні, mls-послідовність, логаріфмічно-ЧМ-сигнал (саме цей сигнал обрано як робочий);
- 3) описано технологію та надано програмне (Matlab) забезпечення для оцінювання ІХ заповненої концертної зали.

Особливості різних джерел звуку показано на Fig. 3-9. Висновок – всі «хлопавки» мають нерівномірний спектр та різну направленість. Перевагою хлопку в долоні є доступність.

Використання додекаедрів із гучномовцями забезпечує всеспрямованість до 2 кГц, проте для повторюваності результатів на 4 кГц та вище треба ретельно встановлювати джерело відносно мікрофона.

Mls-сигнал має рівномірний спектр, проте є чутливим до нелінійних спотворень динаміків, що мають місце на частотах нижче 300 Гц (з’являються тонкі та високі сплески – Fig.7). Виграш від застосування широкосмугових сигналів є частка від ділення довжини вхідного сигналу на довжину вихідного сигналу:

$$B = 10 \lg \frac{T_{ex}}{T_{vix}} = 10 \lg (T_{ex} \cdot \Delta F_{ex}) .$$

Наприклад, для частоти дискретизації 44,1 кГц смуга частот дорівнює 22050 Гц, й для mls-сигналу 2^{16} довжина сигналу є близькою до 1,5 с, тому отримуємо виграш 45 дБ. Щоб взяти результуюче SNR в точці розташування мікрофона, маємо обчислити SNR для сигналу, що випромінюється (від рівня сигналу біля мікрофону віднімаємо рівень шуму біля мікрофону), й додати цей виграш.

Інший недолік mls-послідовності – неузгодженість спектру сигналу із спектром фоновому шуму. Тому пропонується використовувати «червоний» шум (програма із Додатку 2).

На практиці із різних причин реальне SNR буде меншим за розрахункове, що відмічено в статті Бредлі 1996 р. *Одна із причин – низька бітова глибина АЦП та ЦАП. Друга причина – використання сегментування при обчисленнях ДПФ.*

Теоретично можна збільшити SNR шляхом використання більш довгих mls-послідовностей, на що вказував Бредлі. Це вірно для низьких частот, проте не дуже вірно для високих частот, де дається взнаки вплив різних спотворень та змін у часі.

Якщо перед випроміненням робити спектр сигналу червоним (аналогічним спектру фонових завад), а при обробці це спотворення компенсувати, тоді знадобляться надто потужні підсилювач та гучномовець (100 Вт для випроміненої потужності 20 Вт), щоб випромінений сигнал не був спотворений.

Крім того, нелінійні спотворення в гучномовцях можуть призводити не тільки до гострих артефактів-піків, але й до появи фальшивих відбить (Fig.8). Якщо використати mls із червоним спектром, а при обробці компенсувати це спотворення, тоді можна позбутися появи фальшивих відбить (Fig.9), проте рівень фонові завади все ще залишиться підвищеним.

Через змінність параметрів приміщення (конвекція нагрітого повітря від людей, освітлення тощо) на практиці навряд чи вдасться домогтися відношення сигнал-шум вищого за 40 дБ, навіть при застосуванні mls-послідовностей із усередненням та інших вимірювальних методів.

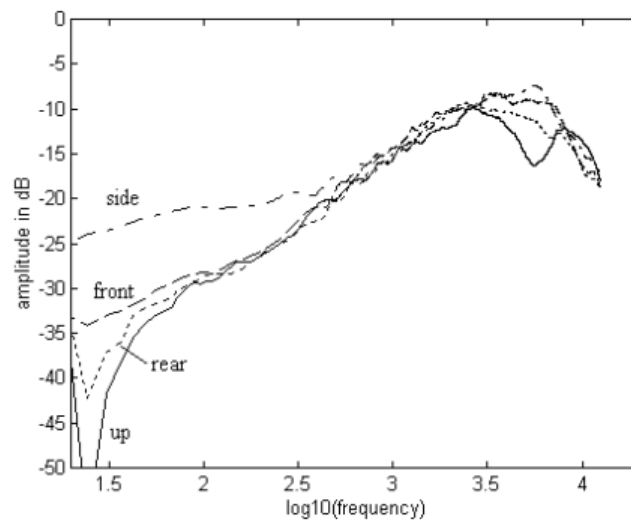


Figure 3: Anechoic spectra from a cap pistol in 4 simultaneous directions. This is a smoothed FFT. On this graph a white spectrum would appear flat. Note the apparant lack of omnidirectionality at LF.

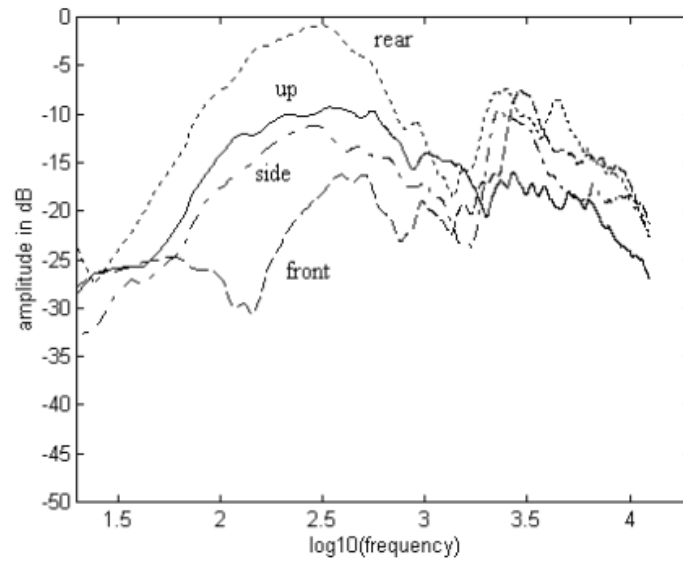


Figure 4: Spectra from a balloon in several directions. Balloon burst with a knife from the rear direction, moving toward the front direction. Smoothed FFT.

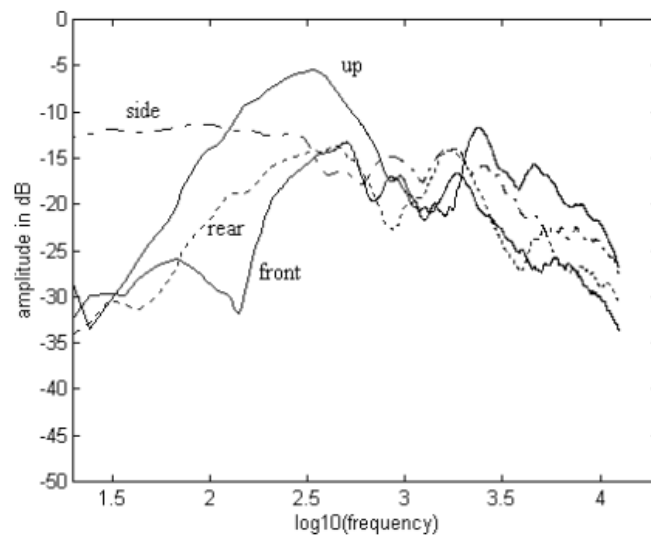


Figure 5: The same as figure 4, but a different balloon. - Note the low frequency directional behavior is quite different.

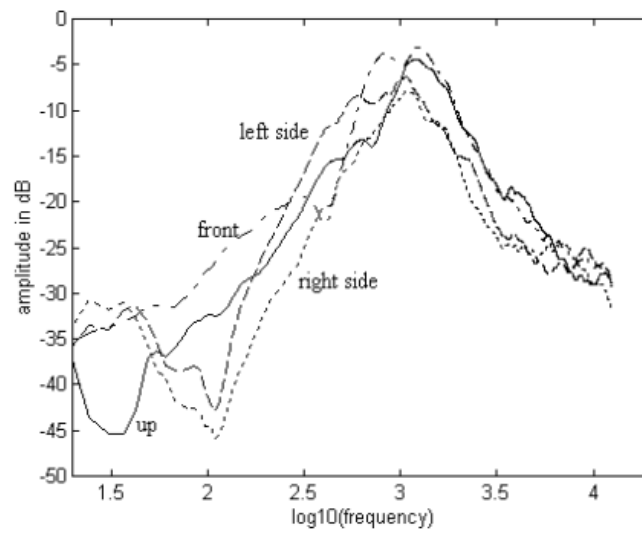


Figure 6: Spectra of a single hand clap in four directions. Smoothed FFT

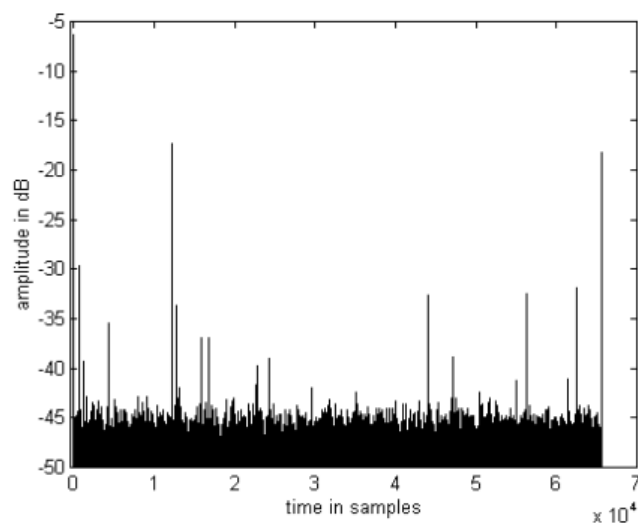


Figure 7: Deconvolution of a 2 second MLS sequence with distortion. Note the numerous spurious reflections and noise.

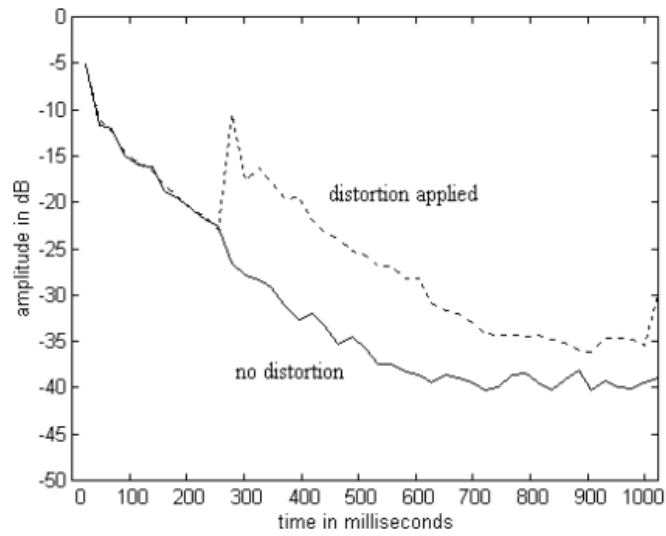


Figure 8: Distorted MLS signal used for the measurement of a reverberator, along with the same reverberator measured with undistorted MLS. 1kHz octave band, integrated in a 43ms window. Note the effect of the distortion is seen mostly in the distortion of the impulse response, not in a decrease in S/N.

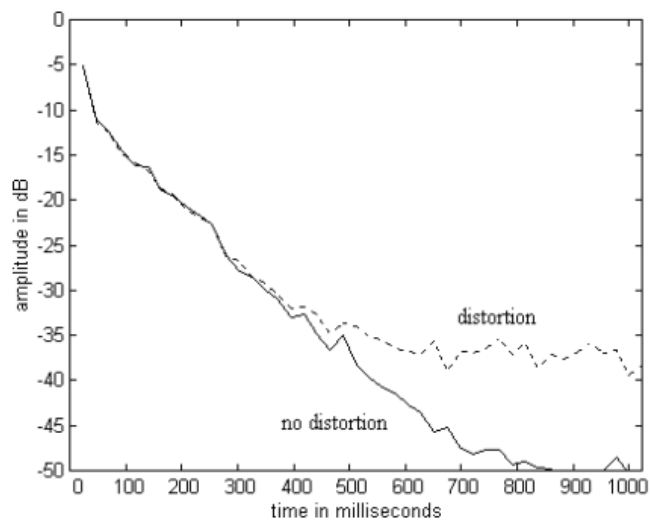


Figure 9: The same reverberator measured with distorted red spectrum MLS.

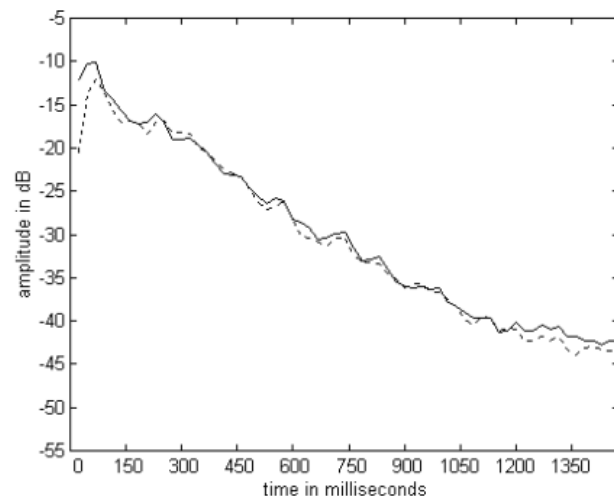


Figure 14: Medial and Lateral - - - response of the hall of figure 1, occupied with the orchestra on stage. Battery operated source hand carried on stage, binaural head connected to a standard cassette recorder in second to last row of the second balcony. 1kHz octave band.

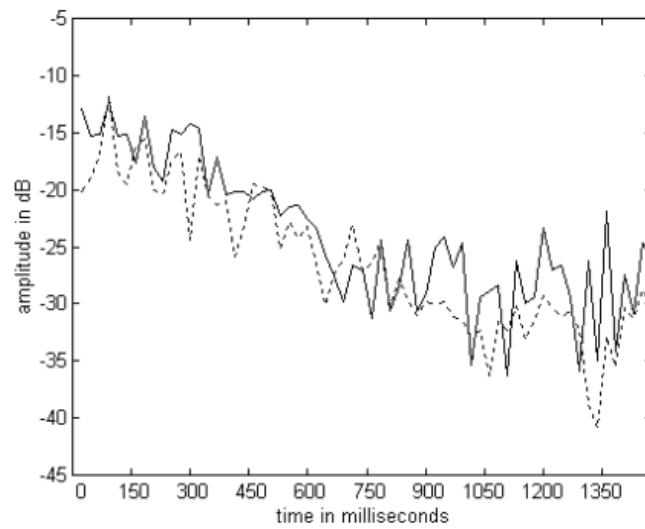


Figure 15: Same measurement as figure 10, but the 125Hz octave band

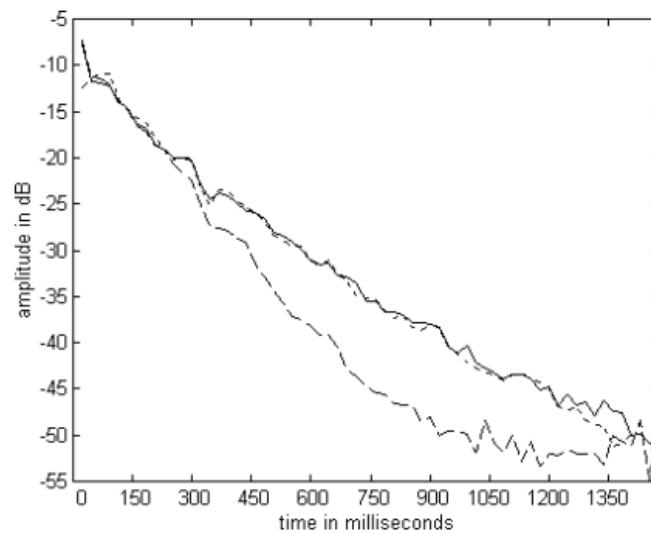


Figure 16: Berlin Staatsoper - source on stage center, microphone in the center of the stalls. = medial, enhancement on. - - - = Lateral, enhancement on. - . - = Medial, enhancement off. 1kHz octave band.

1.4.5. Stan G.-B., Embrechts J.-J., Archambeau D. Comparison of different impulse response measurement techniques. J. Audio Eng. Soc., Vol. 50, No. 4, 2002. P. 249-264.

Виконано порівняння 4 типів тестових сигналів: **MLS** (Maximum Length Sequence – послідовність максимальної довжини), **IRS** (Inverse Repeated Sequence – послідовність зворотного повторювання), імпульси, розтягнуті за часом (**Time-Stretched Pulses**) та частототно-модульовані сигнали (**SineSweep**).

Це порівняння здійснено за допомогою повної, швидкої, надійної та дешевої системи вимірювань. Показано, що за наявності невеликого шуму, MLS та IRS методи здаються більш точними. Навпаки, у тихих умовах метод логарифмічного SineSweep видається найбільш підходящим.

Обґрунтування. У наш час стає все більш очевидною важливість вимірювання бінауральних імпульсних характеристик приміщень з дуже високим співвідношенням сигнал-шум для таких застосувань як віртуальна реальність, ауралізація, просторовість звуків. Після того, як імпульсна характеристика буде точно виміряна, її можна інтегрувати в повний процес ауралізації. Щоб досягти найкращої якості для цього процесу ауралізації, виміряна імпульсна характеристика повинна досягати дуже хорошого співвідношення сигнал-шум (якщо це можливо, більше 80 дБ).

Вимоги до сигналів: 1) повторюваність; 2) максимізація відношення сигнал-шум; 3) стійкість до артефактів у вигляді нелінійних спотворень.

Коротка історія. Техніка MLS вперше була запропонована Schroeder в 1979 році. Незабаром після роботи Шредера (в 1993 р.), був запропонований метод IRS як альтернатива, що дозволяє теоретично зменшити артефакти спотворень, спричинених методом MLS. Через два роки після пропозиції Шредера, Аосіма представив нову ідею для вимірювання імпульсних відгуків, що призвело до техніки імпульсів, розтягнутих за часом. Потім його ідея була розвинута у статті Suzuki, що отримал назву «Оптимальний комп'ютерний

імпульсний сигнал». В 1999 р. Фаріна запропонував логарифмічний спосіб SineSweep, призначений для подолання більшості обмежень, що мають місце в інших методах.

Реалізація.

MLS сигнали. Наша система вимірювань дозволяла генерувати послідовності MLS до 19 порядку (12 секунд для частоти вибірки 44,1 кГц). Перевага MLS-послідовності – некорельованість із фоновою завадою. Недолік – поява артефактів, відомих як «піки спотворення». Боротьба із цими артефактами: 1) використання IRS-послідовностей (артефактові піки зменшуються, проте не зникають – див. рис. нижче); 2) вибір оптимального рівня сигналу, що випромінюється.

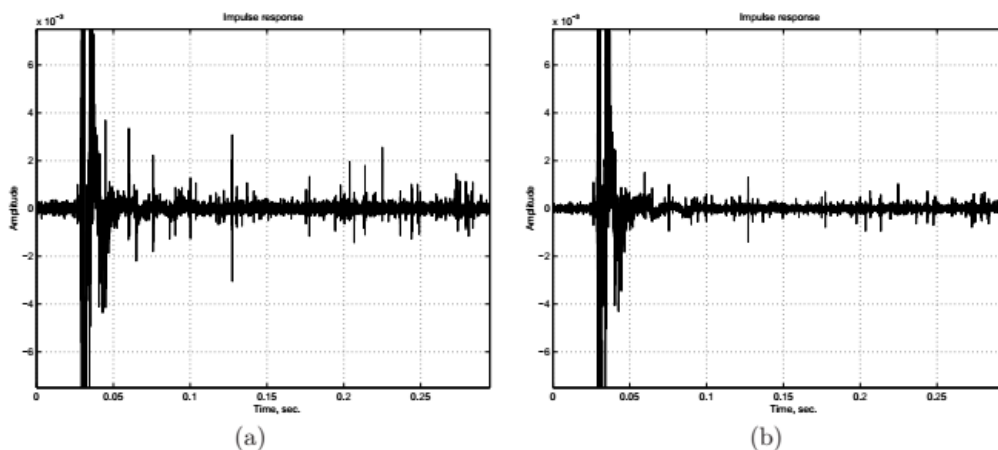


Figure 4: Zoom on the impulse responses obtained via MLS and IRS methods. (a) MLS method. (b) IRS method.

IRS сигнали.

$$x[n] = \begin{cases} mls[n], & \text{if } n \text{ is even, } 0 \leq n < 2L \\ -mls[n], & \text{if } n \text{ is odd, } 0 < n < 2L \end{cases}$$

Довжина IRS сигналів в 2 рази більша за таку для MLS сигналів.

Time-Stretched Pulses. Артефакти у вигляді піків зникають (див. рис. нижче), проте стають видними недоліки (нелінійність) елементів вимірювальної системи, таких, наприклад, як гучномовець.

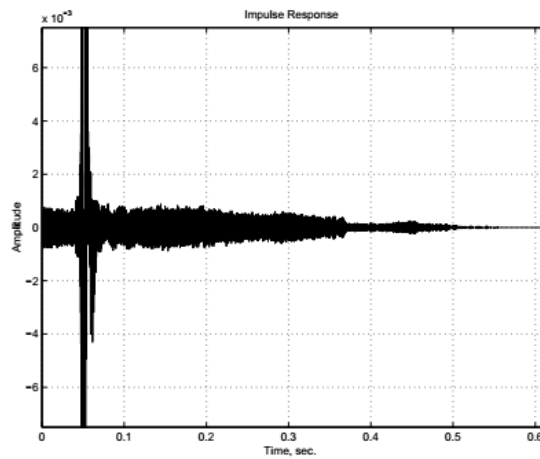


Figure 7: Zoom on the impulse response obtained in an anechoic room when a time-stretched pulse of about 1 second is used. In this case a bad quality loudspeaker has been used in order to emphasize the non linearity of the measurement system.

Logarithmic SineSweep. Методи MLS, IRS та імпульсів, що розтягуються в часі, базуються на припущенні лінійності та часової інваріантності системи, тому їм властиві артефакти, коли ці умови не виконані.

Техніка SineSweep, розроблена Фаріною, вільна від таких обмежень. В її основі лежить наступна ідея: використовуючи експоненціальну швидкість розгортання частоти, можна відокремити імпульсну характеристику системи від імпульсів, спричинених нелінійним спотвореннями. Ці спотворення з'являються перед імпульсним відгуком. Отже: 1) імпульсний відгук, що вимірюється, звільнений від нелінійності; 2) одночасно може бути проведено вимірювання гармонічного спотворення різних порядків.

Аналітично тестовий сигнал SineSweep описується функцією:

$$x(t) = \sin \left[\frac{T\omega_1}{\ln(\frac{\omega_2}{\omega_1})} (e^{\frac{t}{T} \ln(\frac{\omega_2}{\omega_1})} - 1) \right],$$

де ω_1 та ω_2 - початкова та кінцева частоти сигналу довжиною T .

Модель нелінійності. Спричиненої гучномовцем, представлена на рис.8, а результати обробки – на рис.11..

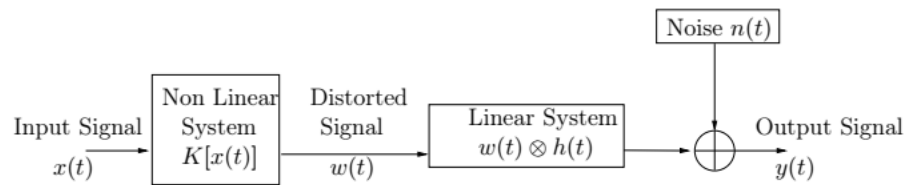


Figure 8: Modelization of the global system including the loudspeaker (considered as a non linear element) and the acoustical space (considered as a perfectly linear system).

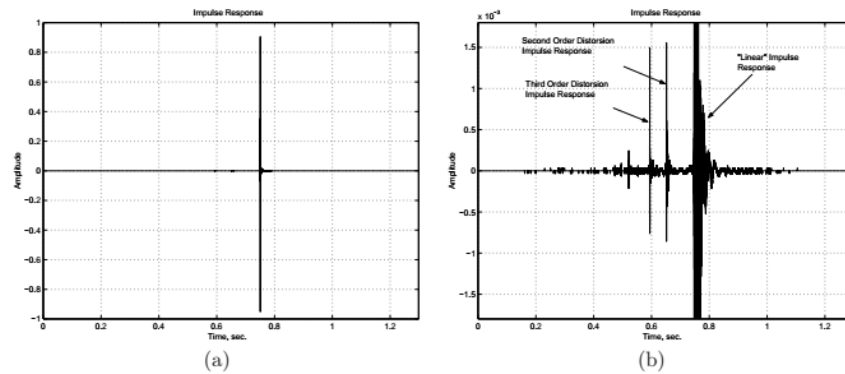


Figure 11: (a) Impulse Response obtained in an anechoic room with a logarithmic SineSweep of 1 second characterized by $\omega_1 = 2\pi 10 \text{ rad/s}$ and $\omega_2 = 2\pi 22000 \text{ rad/s}$. (b) Zoom on this response showing the extraordinary precision of the achievable results.

Експериментальна апаратура. Зазвичай така апаратура є дорогою та громіздкою. В даному випадку програмне забезпечення реалізовано в Matlab 5.3, що здешевлює вимірювальну систему й робить її компактною. Керування випромінювання здійснюється в Matlab за допомогою дуплексної звукової карти, при цьому через цю ж звукову карту одночасно здійснюється запис сигналу, після чого виконується обробка прийнятого сигналу (рис.12). Час вимірювань при цьому складається із часу випромінювання-запису та часу обробки, й становить кілька секунд.

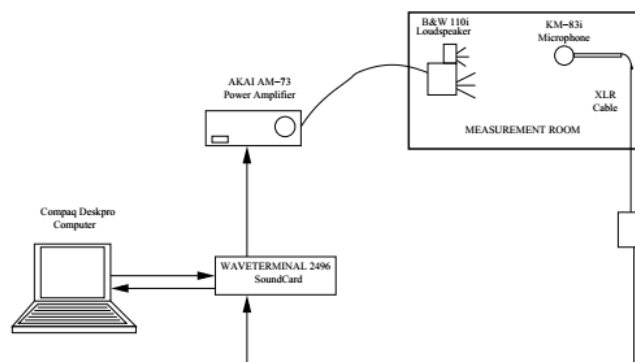


Figure 12: Schematic Representation of the measurement system.

Для корекції нерівномірності АЧХ низькоякісного обладнання, в акустичній кімнаті (рис.14) вимірюють ІХ цього обладнання, після чого коригують сигнал, що випромінюється.

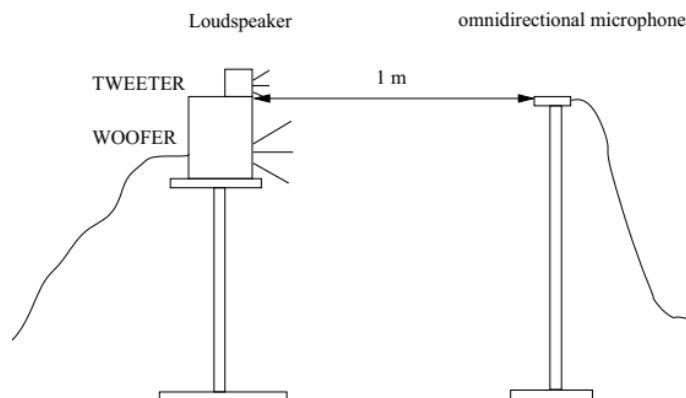


Figure 14: Disposition of the measurement elements in the anechoic room.

Зазначимо, що ми робили трохи інакше – коригували не тестовий сигнал, а результат обробки записаного сигналу. При цьому ми ще й робили придушення високих частот після ділення спектру ІХ на ЧХ підсистеми «гучномовець-мікрофон». Таке придушення треба було б робити і для коригованого тестового сигналу, але автори статті про це мовчать...

Параметри сигналів, що використовувалася. В програмах Матлаб, що були написані, передбачалося, що максимальний порядок MLS та IRS може бути 19.

MLS	IRS	Time-Stretched Pulses	SineSweep
75.5 dB	75.5 dB	83.9 dB	92.5 dB

Table 1: Optimal values of the sound levels at the position of the microphone for each measurement method.

Насправді при дослідженнях було задано порядок MLS та IRS, що становив 16. При частоті дискретизації 44,1 кГц довжина сигналу становила 1,5 с для MLS, та 3 с для IRS сигналу.

Початкова та кінцева частоти SineSweep сигналу становили 10 Гц та 22 кГц, відповідно, а довжина сигналу становила 1,5 с.

Усереднення сигналів не виконувалося.

Рівень шуму в заглушеному приміщенні становив 30 дБ.

Результати вимірювань. У випадку заглушеного приміщення (рис.15) результати MLS та IRS майже не відрізняються, хоча в момент часу приблизно 1,5 с в IRS немає сплеску, який є в MLS.

Відношення сигнал-шум. Оцінювалося як відношення середньої потужності сигналу на виході мікрофона (сигналу на вході кросс-корелятора) до середньої потужності завади в хвості оцінки IX (завади на виході кросс-корелятора). Оцінки SNR для різних тестових сигналів наведено в табл.2.

MLS	IRS	Time-Stretched Pulses	SineSweep
60.5 dB	63.2 dB	77.0 dB	80.1 dB

Table 2: Optimal signal-to-noise ratios for each method.

Приріст 3 дБ для IRS, порівняно із MLS, пояснюється збільшенням вдвічі довжини сигналу. Вражаючий виграш в SNR для Time-Stretched Pulses стає зрозумілим, перш за все, якщо врахувати підвищений рівень сигналу, що випромінювався (на 9 дБ вище, ніж для MLS та IRS). Насправді виграш є меншим й становить приблизно 5 дБ (А.П.). Аналогічна ситуація і для SineSweep сигналу: рівень вхідного сигналу підвищено на 17 дБ, порівняно із таким для MLS та IRS. Якщо це врахувати, то отримаємо $SNR=80-17=63$ дБ, тобто точно такий, як для MLS та IRS. Цікаво, що авторська трактовка дещо відрізняється від нашої, й там дійсно вважається, що SineSweep на 17 дБ виграв в MLS та IRS сигналів. Зрештою, це питання явно спірне... (А.П.).

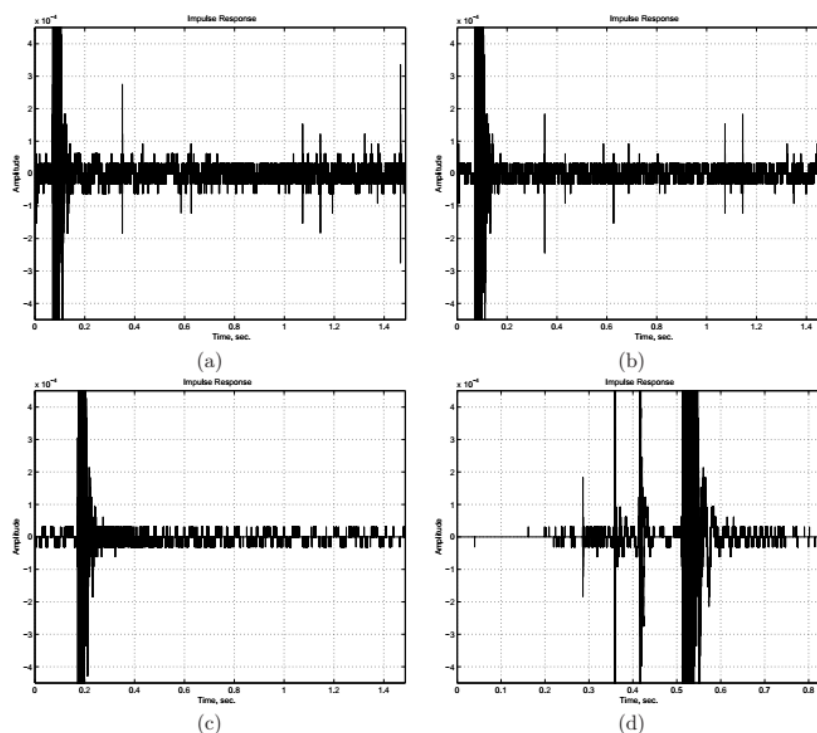


Figure 15: Zoom on the impulse responses obtained in the anechoic room with different methods. (a) MLS (b) IRS (c) Time-Stretched Pulses (d) SineSweep.

Чутливість до імпульсних завад. Вимірювання в заглушеній кімнаті свідчать (рис.16), що MLS (IRS) сигнали значно стійкіші до дії імпульсних завад.

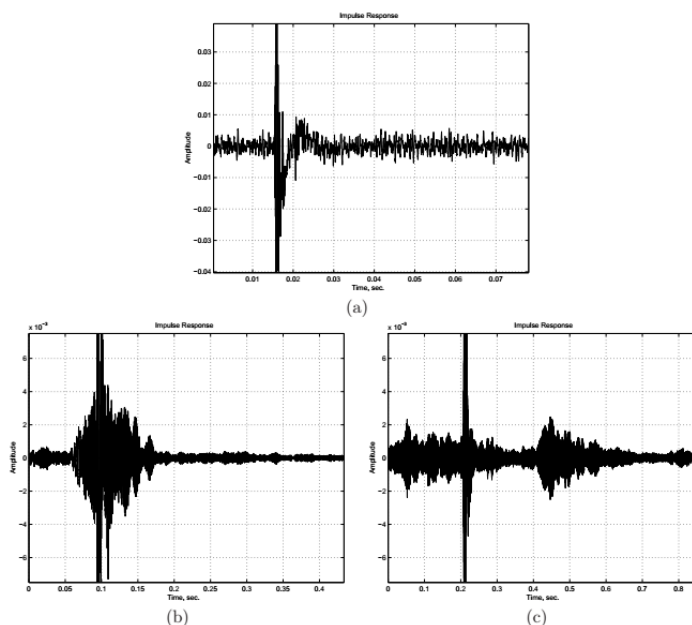


Figure 16: Impulse responses obtained in the anechoic room with different methods when an intense impulsive noise is simultaneously present. (a) MLS (b) Time-Stretched Pulses (c) SineSweep. Notice that the amplitude scales are not identical.

Підсумки:

	Переваги	Недоліки
--	-----------------	-----------------

MLS	Стійкість до шумової та імпульсної завад; невеликий оптимальний рівень сигналу; присмний тембр	Артефакти у вигляді вузьких сплесків через нелінійність обладнання
IRS	Стійкість до шумової та імпульсної завад; невеликий оптимальний рівень сигналу; присмний тембр; менша кількість артефактів через нелінійність обладнання	Артефакти у вигляді вузьких сплесків через нелінійність обладнання менше, але воне є; подвоєна довжина сигналу
Time-Stretched	Артефактів у вигляді вузьких сплесків через нелінійність обладнання немає, але нелінійність спотворює лінійну частину IX	Низька стійкість до шумової завади, що робить необхідним підвищений рівень сигналу; неприємний тембр; не рекомендується для пустих приміщень
SineSweep	Відокремлення проявів нелінійності обладнання від лінійної частини;	Низька стійкість до шумової завади, що робить необхідним підвищений рівень сигналу; неприємний тембр; не рекомендується для пустих приміщень

Література до розділу 1.4

1. M.J.R. Lamothe and J.S. Bradley, "Acoustical characteristics of guns as impulse sources," Can. Acoust. 13(2), 16-24 (1985).
2. Bradley, J. S. Auditorium acoustics measures from pistol shots. Journal of the Acoustical Society of America, 80, 1, pp. 199-205, 1986-07
3. Bradley, J. S. Optimizing the decay range in room acoustics measurements using maximum-length-sequence techniques. Journal of the Audio Engineering Society, 44, (4), pp. 266-273, April 01, 1996
4. D. Griesinger, "Beyond MLS – Occupied hall measurement with FFT techniques," in AES 101th Convention, Preprint 4403, Los Angeles, CA, Nov. 1996.
5. Stan G.-B., Embrechts J.-J., Archambeau D. Comparison of different impulse response measurement techniques. J. Audio Eng. Soc., Vol. 50, No. 4, 2002. P. 249-264.

Лекція 1.5. Штучна голова та її застосування для акустичних вимірювань

Нижче наведено огляд літературних джерел, де виконано порівняння технічних та цінових характеристик макетів голови людини, призначених для вимірювання ІХ приміщень та для запису бінаурального звуку.

1.5.1. M.D. Burkhard and R. M. Sachs Anthropometric manikin for acoustic research. J.Acoust. Soc. Am., Vol.58, No.1, July 1975.

Анотація. Розроблено манекен для слухового апарату і пов'язаного з ним акустичного дослідження з середніми людськими розмірами. Імітація вуха відповідає акустичної реакції анатомічної раковини, вушного каналу і барабанної перетинки, відповідних середнім розмірам, акустичному імпедансу і режимам. Розміри тулуба і голови засновані на опублікованих антропометричних даних, а вушна раковина - на даних, отриманих в ході проведених досліджень. Вушний канал і барабанна перетинка були узгоджені з моделлю Zwislocki. Виявлено, що звуковий тиск на вході вушного каналу слабо залежить від поверхні або опору шкіри голови. Тестові вимірювання показали, що манекен, названий KEMAR, відповідає середньому акустичному відгуку людини у вільному полі.

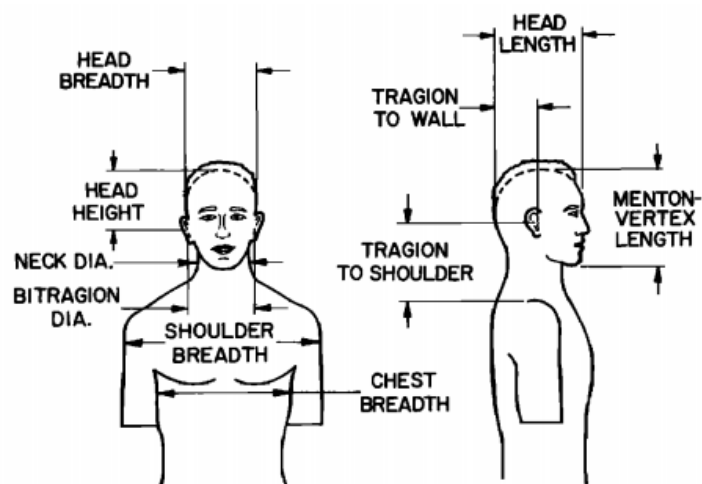


FIG. 1. Anthropometric measures used in design of KEMAR.

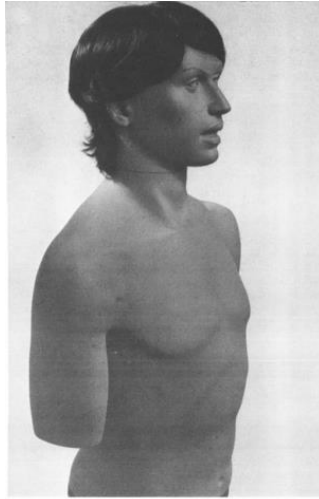


FIG. 5. Photograph of KEMAR with a wig.

TABLE I. Dimensions for KEMAR and average human adults, in centimeters.

	Median male	Median female	Average human	KEMAR
Head breadth	15.5	14.7	15.1	15.2
Head length	19.6	18.0	18.8	19.1
Head height	13.0	13.0	13.0	12.5
Bitracion diameter	14.2	13.5	13.85	14.3
Tragion to wall	10.2	9.4	9.8	9.65
Tragion to shoulder	18.8	16.3	17.55	17.5 ^a
Neck diameter	12.1	10.3	11.2	11.3
Shoulder breadth	45.5	39.9	42.7	44.0
Chest breadth	30.5	27.7	29.1	28.2
Menton vertex length	23.2	21.1	22.15	22.4

^aAdjustable over ± 1.27 cm.

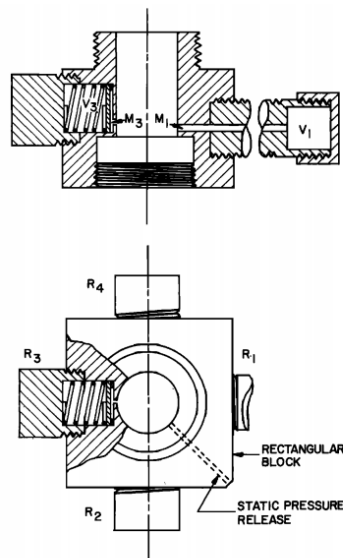


FIG. 2. Schematic drawing of Zwislocki eardrum simulator.

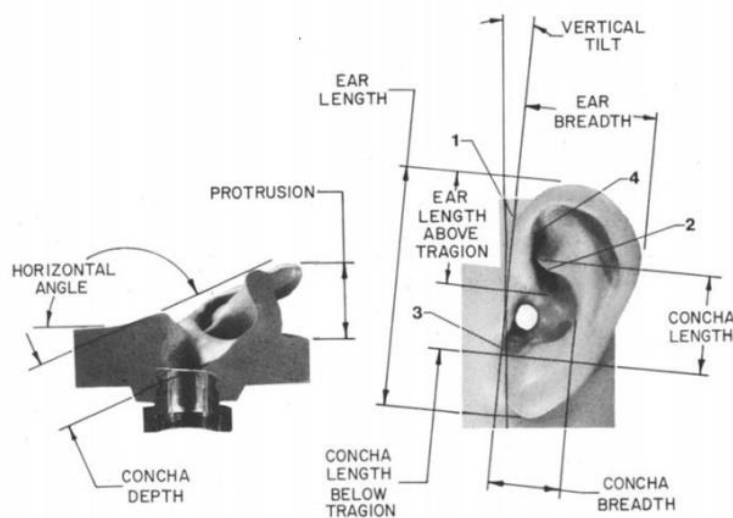


FIG. 3. Auricle measures used in design of KEMAR. (1) Upper pinna-skull notch; (2) antihelix; (3) inter-tragal notch; (4) crus of helix.

TABLE II. External ear dimensions.

Dimension		Averages			Standard deviation			KEMAR	50% Male ^a	50% Female ^a	Average
		12 Male	12 Female	Overall	12 Male	12 Female	Overall				
Ear length	cm	6.85	6.24	6.55	0.59	0.38	0.58	5.89	6.35	5.84	6.10
Ear length above tragon	cm	3.30	3.07	3.19	0.41	0.20	0.34	2.7	3.04		
Ear breadth	cm	3.77	3.36	3.57	0.24	0.27	0.33	3.1	3.55	3.3	3.42
Ear protrusion	cm	2.28	2.03	2.16	0.22	0.23	0.26	1.85	2.10		
Ear protrusion angle	deg	156.7	155.1	155.9	8.6	9.7	9.0	158			
Vertical tilt front view ^b	deg	3.0	2.7	2.9	3.2	3.6	3.1	7			
Vertical tilt side view ^b	deg	7.6	4.7	6.2	2.8	3.4	2.8	6			
Concha volume	cm ³	4.65	3.94	4.30	0.76	0.81	0.85	4.0			
Concha length	cm	2.73	2.53	2.63	0.23	0.20	0.24	2.45			
Concha length, tragon to lower notch	cm	1.74	1.62	1.68	0.16	0.16	0.17	1.82			
Concha breadth	cm	1.88	1.72	1.80	0.21	0.21	0.22	1.57			
Concha breadth tragon to helix	cm	1.82	1.65	1.73	0.27	0.22	0.25	1.39			
Concha depth	cm	1.29	1.29	1.29	0.12	0.08	0.10	1.33			

^aDreyfus (1967).

^bFour males and four females.

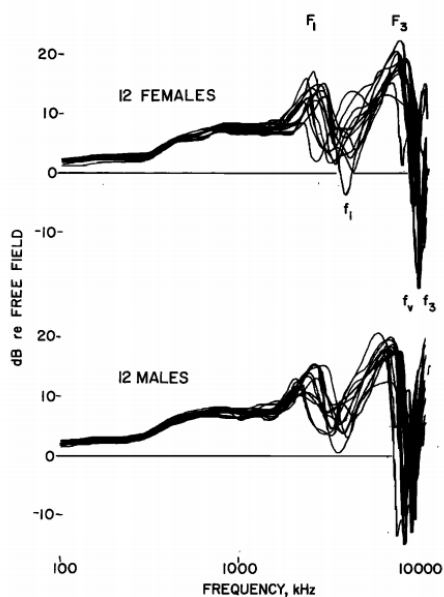


FIG. 4. Ear canal entrance sound pressures, 12 males, 12 females, sound source at 90° azimuth angle, 1m. Free field is the reference pressure condition.

TABLE III. Mean values of frequencies of pressure maxima and minima at the open ear canal entrance for 12 male and 12 female subjects; sound incidence from 90°; loudspeaker source at 1 m. Equivalent lengths based on $c=354$ m/sec for L_1 and L_3 , $c=349.5$ for L_2 and L_4 .

	Mean (kHz)	Std. dev. (kHz)	Equivalent length (mm)
F_1 Male	2.45	0.25	35.6
Female	2.56	0.26	34.1
F_3 Male	7.06	0.57	37.1
Female	7.66	0.58	34.2
f_1 Male	3.63	0.40	24.4
Female	3.69	0.53	24.0
f_3 Male	9.81	0.59	26.1
Female	10.48	0.60	25.4
f_6 Male	8.66	0.59	
Female	9.10	0.51	

*11 males, 11 females.

*10 males, 5 females.

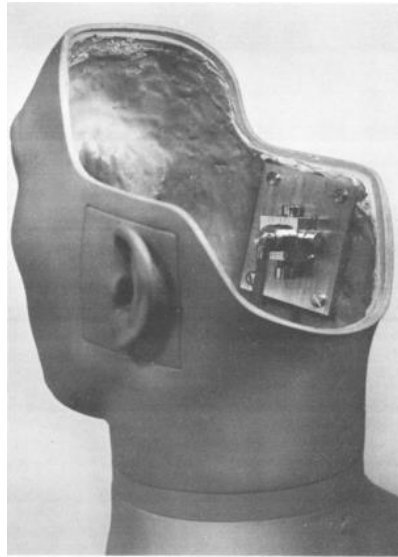


FIG. 6. Photograph of the interior of KEMAR showing eardrum simulator with a Brüel & Kjær 1/8-in. microphone and a right-angle adapter.

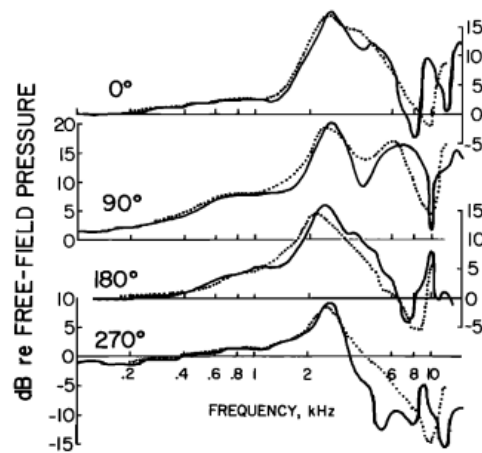


FIG. 7. KEMAR eardrum sound pressure for a free sound field with source azimuth angles of 0°, 90°, 180°, and 270°. —KEMAR; ... Shaw composite.

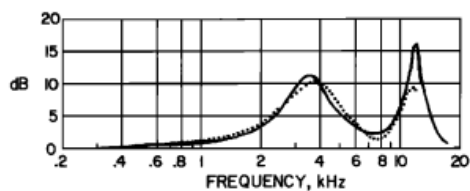


FIG. 8. Ratio of the eardrum to ear canal entrance sound pressure. —KEMAR ...Shaw composite.

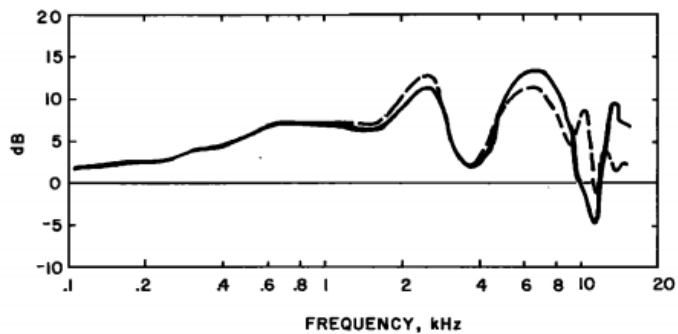


FIG. 9. Sound pressure at the entrance of the open ear canal on KEMAR, both ears, with sound source on the same side as the ear being measured. These curves should be compared to Fig. 4. Reference sound pressure is free field. —right; - - - left.

TABLE IV. Frequencies of pressure maxima and minima at the open ear canal entrance on KEMAR right and left ears. Equivalent length based on $c = 345$ m/sec.

		kHz	Equivalent length (mm)
F_1	right	2.5	34.5
	left	2.5	34.5
F_3	right	6.6	38.0
	left	6.4	39.2
f_1	right	3.7	23.3
	left	3.6	23.9
f_3	right	11.0	23.5
	left	11.3	22.9
f_v	right		
	left	9.0	

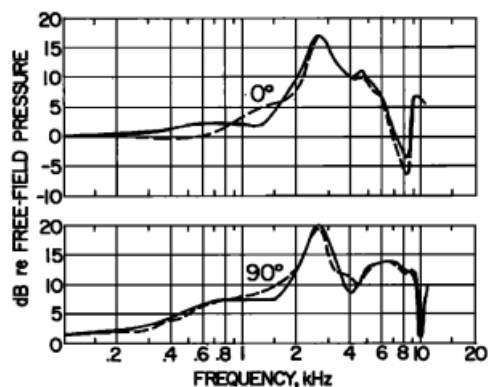


FIG. 10. Eardrum sound pressure in KEMAR showing effect of torso reflections, for sound from front and side. —with torso; - - - no torso.

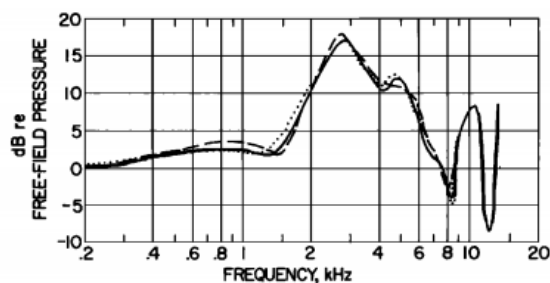


FIG. 11. Eardrum sound pressure in KEMAR showing effect of neck length on torso reflections. ... 18.8 cm; — 17.6 cm; - - - 16.3 cm.

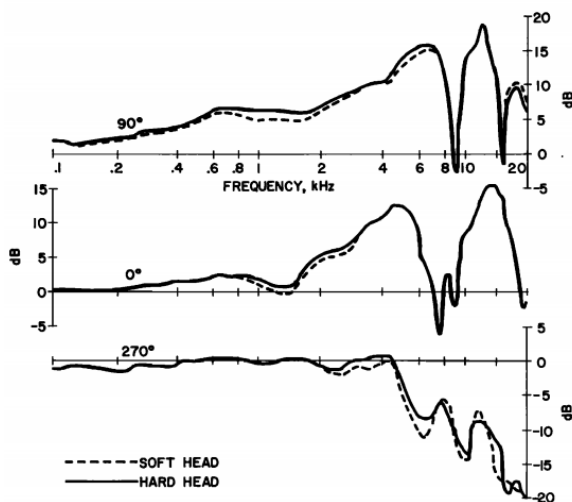


FIG. 12. Comparison of blocked ear entrance sound pressure for a hard fiberglass-reinforced polyester head and a soft sponge RTV head on KEMAR.

Інформація про сучасного виробника KEMAR:

Компанія GRAS Sound & Vibration, заснована в 1994 році піонером датської акустики Гуннаром Расмуссенom, є світовим лідером у галузі звукової та вібраційної промисловості.

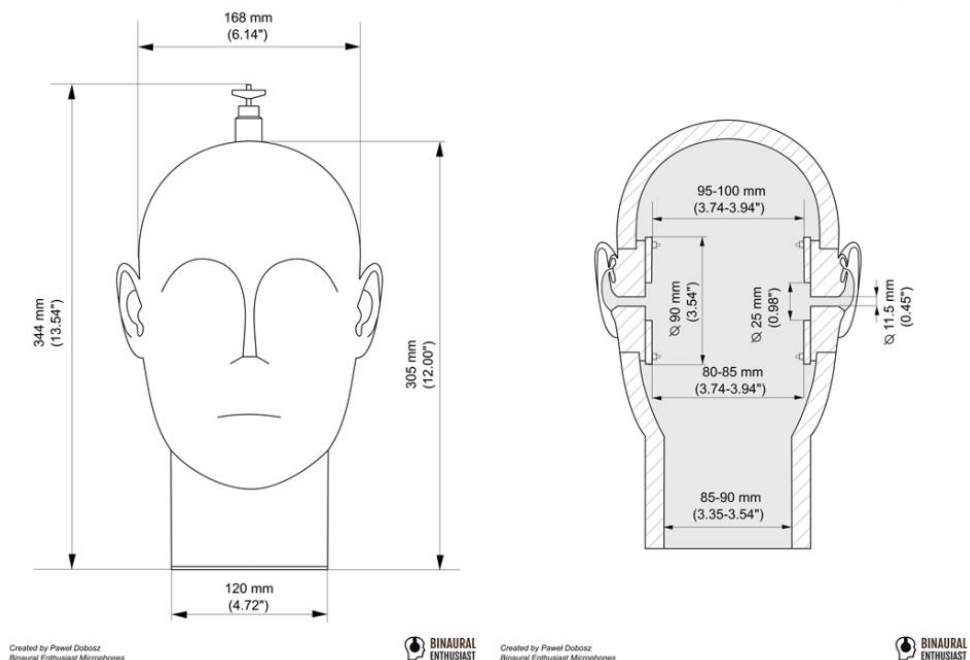
<https://www.grasacoustics.com/products/head-torso-simulators-kemar>

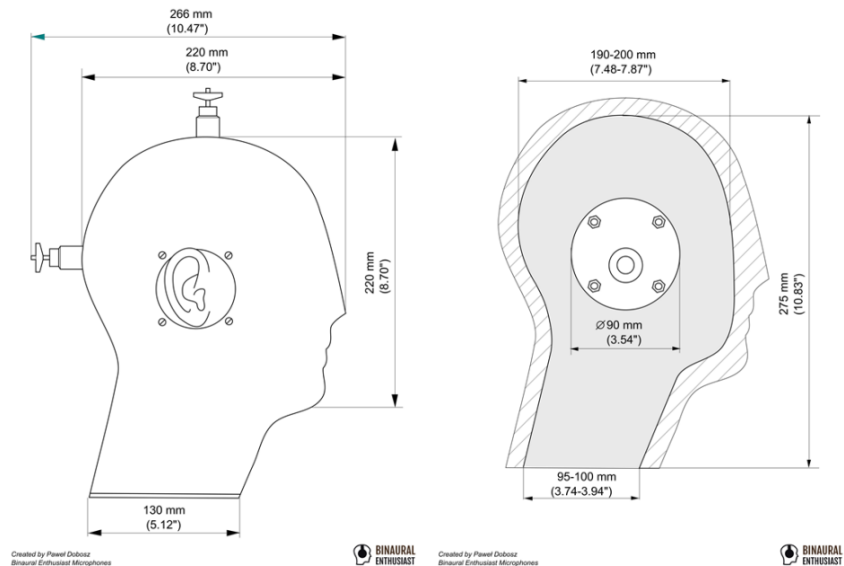
Основна відмінність між KEMAR і стандартними імітаторами голови і торса (HATS) полягає в тому, що KEMAR заснований на великому статистичному дослідженні середнього людського тіла, що означає, що KEMAR HATS має ті ж акустичні властивості, що і пересічна людина.

1.5.2. Інші варіанти виконання штучної голови

1.5.2.1. Польща, Pawel Dobosz (Binaural Enthusiast)

<https://binauralenthusiast.com/product/etiam-ullamcorper-dollor-5/>





1.5.2.2. Amazon. Dummy head and binaural microphone: synthetic head made of MDF

<https://www.amazon.co.uk/Dummy-head-binaural-microphone-built/dp/B07JKXP6TG>



Технічні характеристики:

Розміри виробу 25 x 24 x 29 см; 4,36 кілограма

Сумісні пристрої Mischpult, портативний Aufnahmegeräte

З'єднувач XLR bis XLR

Апаратний інтерфейс 3,5 мм аудіо

Напруга 48 вольт

Ціна: 398-611 €

https://www.thomann.de/gr/soundman_dummy_head_john_black.htm

1.5.2.3. Neumann KU 100 (dummy head) 8.099–8400 €

<https://sennheiserstore.com.ua/ru/ku-100-dummy-head.html>

<https://www.bhphotovideo.com/c/product/146368->

REG/Neumann_KU_100_KU_100_Dummy_Head.html

<https://soniccircus.com/product/neumann-ku-100-dummy-head-stereo-microphone/>



Мікрофон має дві мікрофонні капсули, вмонтовані в «вуха». Діаграма спрямованості подібна вушної, тому при використанні в запису бінауральне сприйняття «переносить» слухача в місце справжньої дії на відміну від інших просторово-пов'язаних методів запису, де саме акустична подія «присувається» до слухача.

Звідси випливає і область застосування даного мікрофона: запис радіопостановок, живих концертів, «круглих столів», а також всіх видів природних звуків.

Крім того, мікрофон широко застосовується в аудіометричних і аналітичних цілях.

Акустичний принцип роботи: Датчик тиску

Діапазон частот: 20 Гц ... 20 кГц

Чутливість при 1 кГц на 1 кОм: 20 мВ / Па

Номінальний опір: 50 Ом збалансований, 200 Ом незбалансований

Номінальний опір навантаження: 1 кОм

Еквівалентний рівень шуму, CCIR1): 29 дБ

Еквівалентний рівень шуму, зважений A1: 16 дБ-А

Відношення сигнал/шум, CCIR1 (відн. 94 дБ SPL): 65 дБ

Співвідношення сигнал/шум, зважене A1 (відн. знач. 94 дБ SPL): 78 дБ

Максимальна SPL для THD 0,5%: 135 дБ

Максимальна SPL для THD 0,5% при попередньому ослабленні: 145 дБ

Максимальна вихідна напруга: 8 дБу

Напруга живлення (P48, IEC 61938): 48 В \pm 4 В/6х1,5 В

Споживання струму (P48, IEC 61938): 2 х 3,5 мА

Відповідний роз'єм: XLR3F / XLR5F

Вага: 3500 г.

Висота: 280 мм, Ширина 180 мм, Глибина 220 мм

1.5.2.4. HMS II.3 (Code 1230) - Head Measurement System with Ear Simulator and Mouth Simulator

https://www.head-acoustics.com/eng/telecom_hms_II_3.htm



Проф. доктор-інж. Клаус
Генуїт, засновник та керуючий
директор HEAD acoustics GmbH та
HEAD Genuit Foundation

1.5.2.4. Манекен голови – розробка кафедри АМЕС

Голова чоловіча чорна матова з пінопласту у велюровій тканині.

Внизу є отвір для кріплення на трубу або підставку.

Вага 100 г.

Висота 28 см.



Застосування: оцінювання бінауральних ІХ приміщення.

Лекція 1.6. Банки бінауральних імпульсних характеристик (ІХ)

Проаналізуємо статті «Do We Need Dereverberation for Hand-Held Telephony?» та «A binaural room impulse response database for the evaluation of dereverberation algorithms» [1, 2]. Мотивація даного аналізу натупна.

Оскільки ми хочемо дослідити вплив ранніх відбить на розбірливість мови, маскованої шумом, нам потрібні бінауральні ІХ приміщень, отримані, по-перше, в приміщеннях певного типу (нас в першу чергу цікавлять лекційні приміщення), а по-друге, для різних відстаней між слухачем і диктором.

Бінауральні ІХ, призначені для слухових апаратів та hands-free телефонів, представлено в базі AIR 1.4, описаній в [1]. Згодом база AIR 1.4 була розширена за рахунок додавання ІХ для звичайної телефонії [2].

Ми хочемо розібратися в наступних речах:

- 1) чи підходять нам ІХ бази AIR 1.4?
- 2) навіщо при записі ІХ для hand-free телефонії використовувалися два мікрофони?

1.6.1. Аналіз статті «Do We Need Dereverberation for Hand-Held Telephony?» [2]

Вплив реверберації зазвичай враховують в hands-free-телефонії і в слухових апаратах. Вважається, що в звичайній телефонії (hand-held) цей вплив несуттєвий. В даній статті показано, що це не завжди так.

В ITU-T G.191 Software Tool Library (STL) є набір ІХ приміщень (RIR), який зазвичай використовують для тестування мовних кодеків. Ці ІХ оцінювалися в офісі і переговорній кімнаті в умовах, коли динамік спрямований на мікрофон. Однак ці умови не відповідають реальним умовам використання звичайного телефону, для яких характерне непрямий напрям випромінювання звуку, а також дифракційні ефекти, зумовлені впливом голови.

У статті [1] описано запропонований набір ІХ, призначений для бінаурального тестування слухових апаратів. Однак в цій статті відсутні ІХ, придатні для звичайної телефонії.

В даній статті переслідується 2 мети. По-перше, описуються нові ІХ, придатні для звичайної телефонії, що дозволяє розширити наявні бази ІХ. По-друге, наводяться результати тестування, об'єктивними і суб'єктивними методами, якості кодеків вузькосмугової, широкосмугової та надширокосмугової телефонії.

1.6.1.1. ITU-T G.191 REVERBERATION MODULE

Характеристики ITU-T G.191 Software Tool Library (STL) представлені в табл. 1.

Room	d_{LM} (m)	RT (s)	DRR (dB)
Office	0.5	0.34	18.64
Office	1.0	0.25	17.40
Conference room	0.5	0.20	16.55

Table 1: Acoustic properties of the STL room impulse responses.

І хоча ревербераційний модуль був доданий до цього інструментарію, все ще залишилося кілька недоліків. Все ІХ були записані без штучної голови, в умовах прямого поширення звуку від гучномовця до мікрофона. Ці умови не відповідають реальним умовам використання звичайного телефону, для яких характерне непрямий напрям випромінювання звуку, а також дифракційні ефекти, зумовлені впливом голови. Крім того, при генеруванні ІХ не використовувалися стандартні положення пристроїв.

Критиці також піддається форма обвідної ІХ з набору STL, приклад якої наведено на рис. 1. У цій обвідній відзначимо три характерних ділянки: 1) спад при переході від прямого звуку до реверберації; 2) спад звуку через

реверберацію; 3) обрив спаду звуку. Останній обрив призводить до завищених значень відношень енергій прямого звуку і реверберації, невластивих як hands-free, так і hand-held телефонії. Як наслідок, використання STL призводить до ненатурального звучання, в порівнянні з таким для ІХ з нормальним спадом рівня звуку. Нарешті, набір STL не враховує затінення звуку головою, а час реверберації є занадто малим.

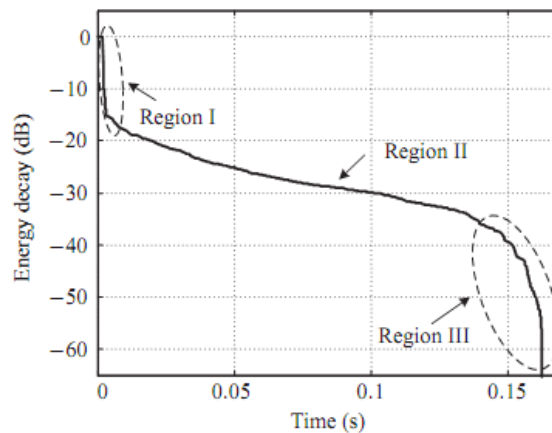


Figure 1: Energy decay curve of STL RIR, measured in an office room at $d_{LM} = 0.5$ m.

1.6.1.2. Процедура вимірювань

Вимірювальна система

Вимірювання проводилися в різних приміщеннях і при різних положеннях телефону. Для вимірювань використовувалася штучна голова зі штучним ротом HMS II.3 фірми HEAD.

Макетом телефону була пластикова коробка розміром 6x12x3 см, в яку було вмонтовано два всеспрямованих вимірювальних мікрофона Beyerdynamic MM1. Другий мікрофон призначений для оцінки можливості використання двоканальних алгоритмів. Обидва мікрофона розміщені на відстані 2 см в передній частині телефону. За допомогою спеціального кріплення макет телефону прикріплювався до штучної голові (рис. 3).



Figure 3: HEAD acoustics HMS II.3 artificial head with the two-microphone mock-up phone clamped in the HEAD acoustics HHP 3 hand-held positioner.

Крім того, одночасно з цим вимірювалися ІХ для hands-free телефонії, для чого телефон містився в відповідне стандартне місце (Hands-Free Reference Point - HFRP). *Стаття [2] відсилає нас, з приводу штучної голови, до стандарту [3], а з приводу прикріплення до неї макета телефону - до стандартів [4,5]. Звідси висновок про корисність ознайомлення з цими стандартами.* Для запису сигналів використовувався ноутбук із зовнішнім звуковою картою і мікрофонним підсилювачем, з частотою дискретизації 48 кГц і бітової глибиною 24 біта. При всіх вимірюваннях звуковий тиск в опорній точці біля рота (mouth reference point - MRP) становив 89 дБ.

Програмне забезпечення

При запису використовувалося програмне забезпечення Matlab r2009b та Windows 7. Синхронізація запису здійснювалася за допомогою RTProc [6] в фоновому режимі Matlab. В якості тестового сигналу використовувався модифікований свип-сигнал [7]. (Наше зауваження: це не дуже зрозуміло, оскільки для вимірювання бінауральних ІХ застосовувалися MLS-послідовності [1]. Питання: чому тут не використовувалися такі ж MLS-послідовності?)

Система передачі звуку

Для реалістичного моделювання телефонного сигналу, спотвореного реверберацією, використовувалося кілька мовних кодеків. У кодеку CELP реверберація пригнічується за допомогою адаптивної фільтрації в мовному

декодері. Однак гарне приглушення реверберації при цьому не може бути здійснено, оскільки фільтрація проводиться в точці прийому. У більшості інших кодеків таку фільтрацію не виконують через її складність.

Схема генерування тестових сигналів:

1) мовні сигнали з бази [8] згортаються з ІХ при частоті дискретизації 48 кГц;

2) отримані реверберовані сигнали передискретизуються, після чого до них застосовується кодування і декодування для 3-х кодеків з наступними значеннями частоти дискретизації, смуги пропускання і бітового потоку:

- Adaptive multi-rate narrowband codec (AMR-NB) [9] $f_s=8\text{kHz}$, 3.4kHz, 12.2 kbit/s;
- Adaptive multi-rate wideband codec (AMR-WB) [10] $f_s=16\text{kHz}$, 7kHz, 23.05 kbit/s;
- Super-wideband (SWB) speech and audio codec [11] $f_s=32\text{kHz}$, 14kHz, 64 kbit/s.

Для простоти, бітова похибка не моделювалася.

1.6.1.3 Об'єктивне оцінювання

Аналіз ІХ

Оцінювалися такі параметри ІХ:

- RT: час реверберації (оцінювалося за рівнями -5 і -30 дБ);
- DRR: ставлення енергії прямого звуку до енергії реверберації;
- ETR: ставлення ранньої енергії до загальної звукової енергії;
- ELR: ставлення енергій ранніх і пізніх відображень (показник ясності).

У табл. 2 представлені результати вимірювань цих параметрів. Бачимо, що час реверберації істотно відрізняється у випадках ННР (hand-held) і HFRP (hands-free) телефонії. Пояснюється це наявністю прямого звуку в разі HFRP і

непрямого звуку у випадку ННР. Зауважимо, що в кухні і офісі ця відмінність є не дуже істотною.

Room	RT (s)		DRR (dB)		ETR (%)		ELR (dB)	
	<i>HHP</i>	<i>HFRP</i>	<i>HHP</i>	<i>HFRP</i>	<i>HHP</i>	<i>HFRP</i>	<i>HHP</i>	<i>HFRP</i>
Office	0.39	0.52	11.79	5.12	99.17	94.83	24.45	15.90
Kitchen	0.42	0.52	11.18	4.62	98.92	93.53	22.92	15.06
Corridor	1.25	1.47	10.98	4.35	97.56	89.06	17.49	10.45
Stairway	0.86	1.23	12.28	6.41	98.91	93.68	21.69	13.90

Table 2: Channel-based measures calculated directly from the impulse responses. The results are averaged over both channels.

Що стосується DRR, відмінність є істотною у всіх випадках. Порівнюючи табл. 1 і 2, бачимо, що в разі STL час реверберації помітно менше, а DRR помітно більше, що забезпечує хорошу розбірливість мови для IX STL. На рис. 5 показана залежність RT від частоти: для кухні та офісу вона слабка, тоді як для сходового прольоту і коридору вона є помітно вираженою.

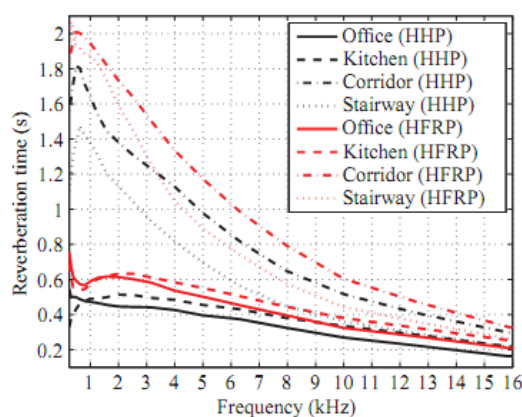


Figure 5: Subband reverberation time (SRT) calculated by means of the Schroeder method in 1/3-octave subbands.

Аналіз когерентності

У приміщеннях з вираженою реверберацією повинен бути вираженим і дифузний характер звукового поля, для відстаней між гучномовцем і мікрофоном >1 м. На рис. 6 показані функції когерентності для офісу і сходового прольоту. З цього малюнка виходить, що якщо для придушення реверберації використовуються алгоритми, що базуються на врахуванні когерентності, хороші результати можна отримати тільки для високих частот, де когерентність низька через малу відстань між мікрофонами.

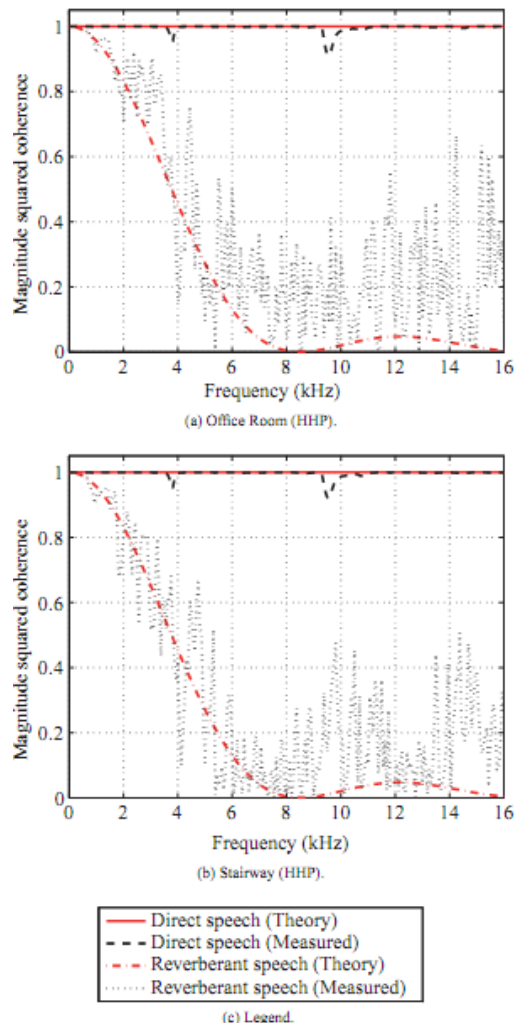


Figure 6: Magnitude squared coherence (MSC) of direct and reverberant speech components for (a) the office in HHP position and (b) the stairway in HHP position.

Вимірювання якості сигналів

У даній роботі такі вимірювання зроблені з використанням міри PEMO-Q [12,13], яка добре враховує вплив реверберації на якість мовних сигналів. Результати наведені на рис. 7.

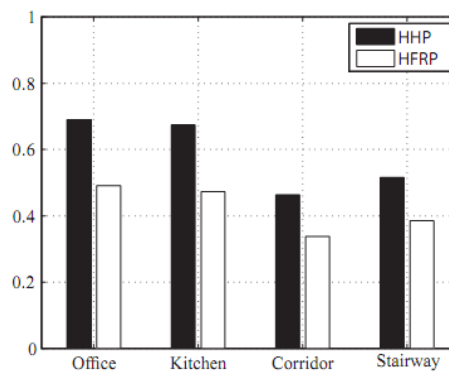


Figure 7: Objective evaluation of the reverberant speech signal using the non-intrusive PEMO-Q measure. A PEMO-Q score of 1 indicates the highest perceptual similarity.

1.6.1.4. Суб'єктивне оцінювання

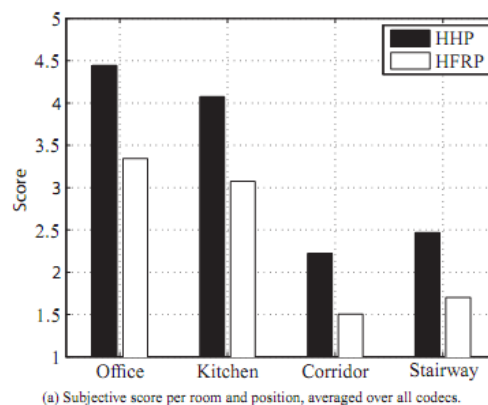
Тестування прослуховуванням

Таке прослуховування виконувалося в заглушеному приміщенні, що знижує рівень зовнішнього шуму на 42 дБ. Для мінімізації спотворень звуку, використовувалися цифровий еквалайзер HEAD Acoustics PEQ V в комбінації з навушниками Sennheiser HD600. Під час тесту з 30 досвідченими слухачами (нормальний слух, вік 24-33 роки) учасникам були представлені 24 сигнали. Чистий мовний сигнал тривалістю 18 секунд оброблявся відповідно до модельованого тракту передачі. Кожен з 8 реверберованих сигналів (4x ННР, 4x HFRP) кодувався кодеком NB, WB і SWB. Для кожної ситуації слухачам було запропоновано оцінити деградацію відповідно до шкали деградації ITU-R BS.1284-1 (див. Табл. 3 [14]).

5.0	Imperceptible
4.0	Perceptible but not annoying
3.0	Slightly annoying
2.0	Annoying
1.0	Very annoying

Table 3: ITU-R BS.1284-1 five-grade impairment scale [19].

Слід зазначити, що слухачів не просили оцінити загальну якість мовлення, а тільки погіршення через реверберацію кімнати. Тому результати, отримані для різних приміщень (рис. 8а) і кодеків (рис. 8б), не є рейтингом якості сигналів.



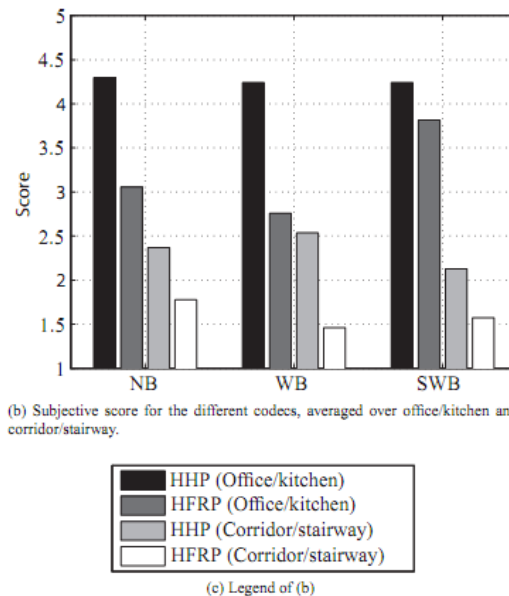


Figure 8: Results of the listening test according to the ITU-R five-grade impairment scale: (a) averaged over the different codecs and (b) averaged over two sets of rooms (office/kitchen and corridor/stairway).

Отримані результати показують, що ступінь деградації практично не залежить від кодеків, що розглядалися.

1.6.1.4. Висновки

Отримані в рамках проведених досліджень ІХ більш реалістичні, ніж ІХ, представлені в базі STL. Показано, що в умовах сильної реверберації (коридор, сходові клітки) доцільне застосування алгоритмів дереверберації як для hands-free, так і для hand-held телефонії.

1.6.2. Аналіз статті «A binaural room impulse response database for the evaluation of dereverberation algorithms»

У даній статті описана база бінауральних ІХ, отриманих з використанням штучної голови для таких приміщень як заглушена кімната, офіс, переговорна кімната і лекційна аудиторія. Основна область застосування цієї бази - використання при тестуванні алгоритмів дереверберації в таких пристроях як слухові апарати і hands-free телефонія. Обґрунтовано перевагу використання бінауральних вимірювань замість монауральних. Для демонстрації можливостей отриманої бази ІХ, наведено приклад придушення реверберації за допомогою

алгоритму, що використовує функцію когерентності. Отримана база AIR представлена для відкритого використання в Інтернеті.

Алгоритми дереверберації можуть широко застосовуватися і вже застосовуються в слухових апаратах і в hands-free телефонії. Для тестування алгоритмів дереверберації потрібно мати у своєму розпорядженні ІХ приміщень. Такі ІХ синтезують, виходячи з параметрів приміщень, або вимірюють в реальних приміщеннях. Однак існуючі ІХ можна застосувати для слухових апаратів, оскільки вони не враховують затінення головою. У даній статті представляється вельми повна база вимірюваних бінауральних ІХ. Виміри проводилися як зі штучною головою, так і без неї, в заглушеному приміщенні, офісі, переговорному приміщенні і в аудиторії.

1.6.2.1. ІХ приміщень

Маючи в своєму розпорядженні ІХ приміщень, можна моделювати звук в приміщенні шляхом згортки чистого сигналу з ІХ приміщення. Недоліком синтезу ІХ методом уявних джерел є обмеженість кількості відбивачів, що ускладнює забезпечення експоненціального характеру загасання енергії. Цього недоліку позбавлений спосіб вимірювання ІХ в реальних приміщеннях, що реалізується з використанням псевдовипадкових послідовностей (MLS-послідовності) або послідовностей PSEQs. На жаль, і тут є одна проблема - вимірювану ІХ не можна змінювати. Це обмеження можна обійти, створюючи великі бази даних, що характеризують різні приміщення.

Нещодавно запропонована база, іменована MARDY (Multichannel Acoustic Reverberation Database at York) [15], непридатна для використання в бінауральних алгоритмах, оскільки запис проводився на 8 мікрофонів, віддалених на 5 см, при цьому не враховувався ефект затінення звуку головою.

Слід зазначити ще такий варіант, де замість процедури згортки чистого сигналу з ІХ використовують запис звуку в реальних приміщеннях. Перевага такого підходу - немає необхідності вважати приміщення лінійною системою, що

в дійсності є спрощенням дійсного стану речей. Однак припущення лінійності системи, що дозволяє використовувати досить просту процедуру згортки для моделювання реверберірованих сигналів, забезпечує значну гнучкість при дослідженнях. Тому підхід, який базується на припущенні лінійності і використанні співвідношення згортки, слід вважати кращим.

1.6.2.2. Бінауральні ІХ приміщень

Далі буде показано, що використання бінауральних ІХ є кращим при тестуванні алгоритмів дереверберації в слухових апаратах. На жаль, в ряді публікацій використовують стереозаписи, а не бінауральні, під якими мають на увазі записи, зроблені з використанням штучної голови. Мікрофони при цьому розташовуються поблизу вушних раковин.

При використанні алгоритмів дереверберації, заснованих на використанні когерентності, важливо, щоб когерентність сигналів на виходах 2-х мікрофонів була низькою в дифузному полі. На рис. 1 показані функції когерентності для різних відстаней між мікрофонами.

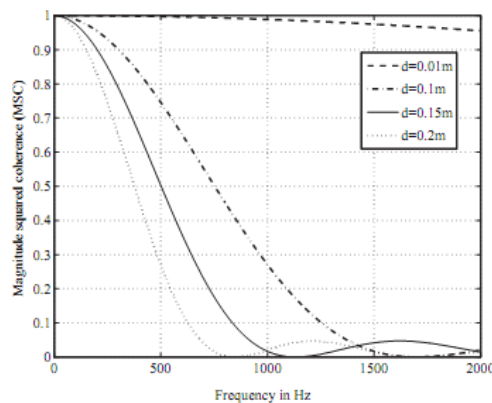


Fig. 1. Theoretical coherence function of two closely spaced microphones for different distances.

Бачимо, що якщо між мікрофонами відстань відповідає відстані від лівого до правого вуха, тобто приблизно $d=0,15$ м, функція когерентності має значення $<0,2$ для частот вище 700 Гц.

Неважко зробити висновок, що алгоритм дереверберації, заснований на використанні когерентності, повинен бути виключений з розгляду, якщо конструкція технічного пристрою не дозволяє зробити відстань між мікрофонами досить великим. Однак даний висновок не зовсім вірний, оскільки не враховується можлива екранування мікрофонів, як це відбувається зі слуховими апаратами, де голова екранує один мікрофон від іншого (рис. 2).

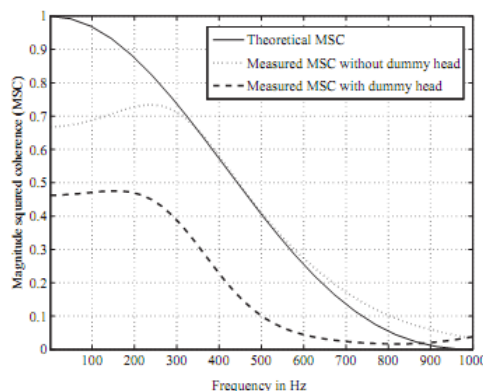


Fig. 2. Influence of the presence of a dummy head on the magnitude squared coherence. Solid line: theoretical MSC using Eq.(2) with $d = 17$ cm. Dashed line: Estimated MSC with a dummy head. Dotted line: Estimated MSC without a dummy head. The two microphones are positioned close to the pinna at a distance of $d = 17$ cm.

Таким чином, робимо висновок, що вимірювання бінауральних ІХ є завданням актуальним.

1.6.2.3. Вимірювальна система

Вимірювальне обладнання

Всі вимірювання проводилися з частотою дискретизації 48 кГц і 24-битами. Для бінауральних записів використовувалася штучна голова HMS2 фірми HEAD. Два мікрофона розташовувалися поблизу вушних раковин на відстані 1 см від вушного каналу. Вимірювання повторювалися після видалення штучної голови, щоб дослідити вплив голови.

MLS-послідовності

Для вимірювань використовувалися MLS-послідовності ступеня 15, що дає довжину послідовності $2^{15} - 1 = 32767$ вибірок або максимальну довжину

імпульсних відгуків 0,68 с (для аудиторії використовувалася ступінь 16, яка веде до максимальної довжини 1,36 с). Сигнали повторювалися 17 раз, і імпульсна характеристика потім обчислювалася як середнє останніх 16 повторень, що призводило до збільшення SNR на $10\lg 16 = 12$ дБ.

1.6.2.4. Характеристики приміщень

Вимірювання проводилися в 4-х приміщеннях. Характеристики одного з приміщень - аудиторії - наведені в табл. 1. Пізніше були додані результати для $d=2,25$ м.

Lecture room
10.80 m x 10.90 m x 3.15 m
1.2 m
1.2 m
0.17 m
4.0 m, 5.56 m, 7.1 m, 8.68 m, 10.2 m
3x glass windows, 1x concrete wall
Parquet
Wooden tables, chairs

Зауваження: в програмі load_air.m, призначеної для роботи з базою AIR, знаходимо правило формування кодів в іменах файлів бази:

- в 1-й позиції - airpar.channel ('0': right; '1': left);
- у 2-й позиції - airpar.head ('0': no dummy head; '1': with dummy head);
- в 3-й позиції - airpar.rir_no (2.25m, 4m, 5.56m, 7.1m, 8.68m, 10.2m)

Приклад:

air_binaural_lecture_0_0_1.mat – правий канал; без голови; $d=2,25$ м

1.6.2.5. Постобробка

Файли бази даних зберігаються у вигляді MAT-файлів з подвійною точністю і 24-розрядних WAV-файлів. Затримка звуку для самої короткої

дистанції ліквідована, проте додаткові затримки звуку для інших дистанцій збережені.

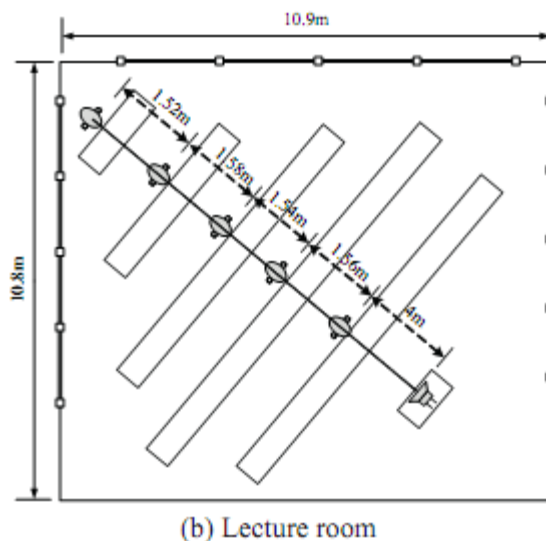


Рис. 3. Схема вимірювань в аудиторії

1.6.2.6. Оцінювання

Час реверберації. Зауваження - наведено дані лише для аудиторії (табл. 2, рис. 4).

Lecture room ($\overline{RT}_{60} = 0.78$ s)	2.25	0.70
	4.00	0.72
	5.56	0.79
	7.10	0.80
	8.68	0.81
	10.2	0.83

Table 2. Overview of the measured reverberation times for each room depending on the source-microphone distance. All values are averages between left and right channel. The overall mean values RT_{60} are given by the average among all distances.

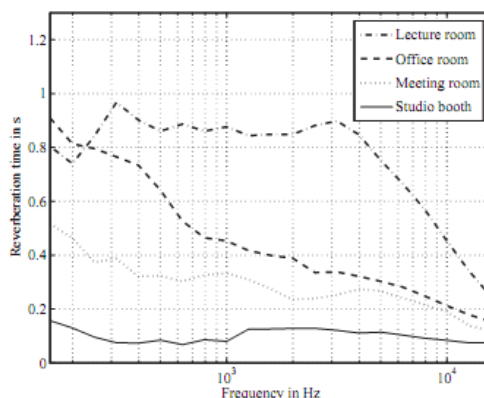


Fig. 4. Subband reverberation time obtained by a 1/3-octave filterbank and the Schroeder method [11]. The values are averaged over all measured positions in the room including averaging of left and right channel.

Приклад застосування. Використовувався мультімікрофонний алгоритм придушення реверберації Аллена [16]. Результати наведені в табл. 3.

Для оцінки використовувалися неінтрузивні міри у вигляді відношення енергій вихідної і реверберованої мови (SRMR) [17] і сегментне відношення сигнал-реверберація (SegSRR). Перед оцінюванням періоди мовчання були видалені за допомогою детектора голосової активності (VAD) мовного кодека AMR-WB [18]. Експеримент проводився на 100 мовних файлах бази даних TIMIT для $f_s = 16$ кГц.

	Δ SRMR	Δ SegSRR
No dummy head	+1.34	+0.67 dB
Dummy head	+2.16	+1.39 dB

Table 3. Reverberation reduction by the dual-channel Allen algorithm with and without the presence of a dummy head in the lecture room. The values indicate the improvement compared to the reverberant speech.

З табл. 3 випливає, що ефективність придушення реверберації вище для записів IX з використанням штучної голови.

1.6.3. Висновки за результатами розгляду 2-х публікацій

1. Автори бази AIR на передній план виставляють той факт, що записи, зроблені за допомогою штучної голови, виправдовують використання мультімікрофонного алгоритму дереверберації як в слухових апаратах, так і в гучномовній (hands-free) телефонії. Проте якщо ставиться інша мета - з'ясування ступеня впливу ранніх відбиттів на розбірливість мови - база AIR цілком підходить. По-перше, вона позбавляє від необхідності змушувати слухачів працювати в реальних приміщеннях. При цьому ясно, що якщо слухачі працюють в реальних приміщеннях поодиночі, це дозволяє створити однакові умови прослуховування, але значно збільшує час випробувань. Якщо слухачі працюють в реальних приміщеннях групою, час випробувань скорочується, але тоді умови випробувань стають неоднаковими для різних слухачів. По-друге, цікаво, наскільки важливі вимірювання IX зі штучною головою для задачі

оцінювання розбірливості мови в аудиторіях та інших приміщеннях. Якщо важливі - актуальною тоді актуальним стає завдання придбання або виготовлення штучної голови.

2. Дуже корисні вказівки авторів роботи [2] щодо організації обробки сигналів, а саме: згортання чистого сигналу та IX , що виконується при частоті дискретизації 48 кГц (щоб не псувати записи IX , зроблені при $f_s = 48$ кГц). Лише після цього реверберовані сигнали мають піддаватися передискретизації (за потреби).

Література до підрозділу 1.6

1. M. Jeub, M. Schafer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in Proc. Int. Conference on Digital Signal Processing (DSP), Santorini, Greece, 2009.

2. M. Jeub et al., "Do We Need Dereverberation for Hand-Held Telephony?" in Proceedings of 20th International Congress on Acoustics, ICA 2010, 23–27 August 2010, Sydney, Australia.

3. ITU-T P.581 (05/2000) SERIES P: TELEPHONE TRANSMISSION QUALITY. TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS. Objective measuring apparatus. Use of head and torso simulator (HATS) for hands-free terminal testing.

4. ITU-T P.64 (11/2007) SERIES P: TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS. Objective electro-acoustical measurements. Determination of sensitivity/frequency characteristics of local telephone systems

5. ITU-T P.340 (08/96) SERIES P: TELEPHONE TRANSMISSION QUALITY. SUBSCRIBERS' LINES AND SETS. Transmission characteristics of hands free telephones.

6. H. Krüger and P. Vary, "RTPROC: A system for rapid real-time prototyping in audio signal processing," in Proc. 12th IEEE/ACM International

Symposium on Distributed Simulation and Real-Time Applications DS-RT, Vancouver, Canada, 2008, pp. 311–314.

7. A. Telle, C. Antweiler, and P. Vary, “Der perfekte Sweep - Ein neues Anregungssignal zur adaptiven Systemidentifikation zeitvarianter akustischer Systeme,” in Proc. of German Annual Conference on Acoustics DAGA, Berlin, Germany, 2010.

8. P. Kabal, “TSP speech database,” Tech. Rep., Department of Electrical & Computer Engineering, McGill University, Montreal, Quebec, Canada, 2002.

9. 3GPP TS 26.071, Adaptive Multi-Rate (AMR) speech codec; General description, 3GPP, 2004.

10. 3GPP TS 26.171, Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description, 3GPP, 2004.

11. B. Geiser, H. Krüger, H.W. Löllmann, P. Vary, D. Zhang, H. Wan, H. T. Li, and L. B. Zhang, “Candidate proposal for ITU-T super-wideband speech and audio coding,” in Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP), Taipei, Taiwan, 2009, pp. 4121–4124.

12. S. Goetze, E. Albertin, M. Kallinger, A. Mertins, and K.-D. Kammeyer, “Quality assessment for listening-room compensation algorithms,” in Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP), Dallas, TX, USA, 2010.

13. R. Huber and B. Kollmeier, “PEMO-Q—a new method for objective audio quality assessment using a model of auditory perception,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 6, pp. 1902–1911, 2006.

14. ITU-R Rec. BS.1284-1, General methods for the subjective assessment of sound quality, ITU, 2003.

15. J.Y.C. Wen, N.D. Gaubitch, E.A.P. Habets, T. Myatt, and P.A. Naylor, “Evaluation of speech dereverberation algorithms using theMARDY database,” in Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC), Paris, France, 2006.

16. J.B. Allen, D.A. Berkley, and J. Blauert, "Multimicrophone signal-processing technique to remove room reverberation from speech signals," J. Acoust. Soc. Am., 1977.
17. T.H. Falk and W.-Y. Chan, "A non-intrusive quality measure of dereverberated speech," in Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC), Seattle, USA, 2008.
18. TS 26.194, Adaptive Multi-Rate - Wideband speech codec, Voice Activity Detector, V6.0.0, 3GPP, 2004.

Лекція 1.7. Акустична експертиза приміщень для мовленнєвих презентацій

Така експертиза може виконуватися двома шляхами:

- 1) суб'єктивне оцінювання [1];
- 2) об'єктивне оцінювання [2].

1.7.1. Суб'єктивне оцінювання розбірливості мови артикуляційним методом

Останнім часом, з огляду на економію часових, фінансових та людських ресурсів, значну увагу приділяють об'єктивному підходу, який дозволяє автоматизувати процедуру вимірювання. Проте, незважаючи на ресурсоемкість суб'єктивного підходу, його цінність залишається безсумнівною, оскільки, по-перше, такий підхід призводить до результатів, зрозумілих будь-якій людині. По-друге, одержані результати є еталонними в тому сенсі, що в подальшому можуть бути використані для калібрування об'єктивних систем вимірювання.

Тому на сьогодні актуальним є завдання максимального полегшення праці людей, які приймають участь у вимірюваннях розбірливості мови шляхом артикуляційних випробувань. Зробити це можна шляхом автоматизації окремих етапів артикуляційних випробувань, із використанням комп'ютерної техніки та відповідного програмного забезпечення для запису й відтворення мовлення дикторів, моделювання ліній зв'язку та керування процедурою (рис. 1) артикуляційних випробувань.

Російський стандарт ГОСТ Р 50840-95 припускає використання комп'ютерної техніки та регламентує проведення артикуляційних випробувань за схемою рис. 1б. На жаль, українські інженери не можуть скористатися цим стандартом за відсутністю відповідного програмного забезпечення й тому змушені користуватися застарілим стандартом СРСР ГОСТ 16600-72.

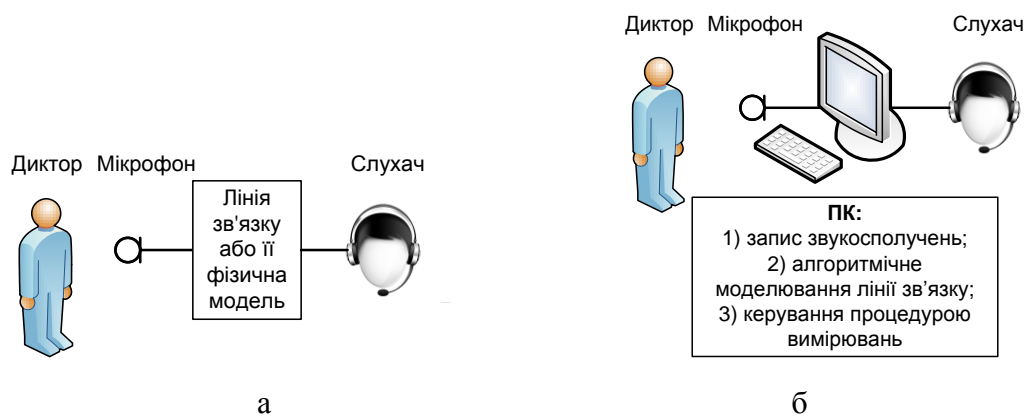


Рис. 1. Традиційна схема артикуляційних випробувань (а) та схема із застосуванням ПК (б)

Важливим кроком на шляху до створення вітчизняного державного стандарту, що відповідає сучасним вимогам, можна вважати роботи [3], [4], де запропоновано та випробувано десять артикуляційних таблиць українських слів. Одним із недоліків вказаних праць є відсутність пропозицій щодо автоматизації процедури артикуляційних випробувань. Іншим недоліком є використання таблиць слів, оскільки відомо, що використання таблиць односкладових звукосполучень забезпечує вищу достовірність випробувань [5]. Зазначимо також, що не досить коректним є прослуховування слів, спотворених шумом, через комп'ютерні колонки, оскільки на одержані таким чином результати можуть суттєво впливати особливості приміщення, в якому відбуваються випробування.

До недоліків праці [4] слід віднести й фактичну невизначеність характеру забарвленості шуму, одержаного за допомогою генератора ANG-2200 [6]. Між тим, питання забарвленості маскувального шуму є важливим, оскільки до сьогодні не існує чіткої відповіді стосовно маскувальних можливостей білого шуму, порівняно із рожевим та коричневим шумами, при малих відношеннях сигнал-шум. В [7] зроблено спробу дослідити це питання шляхом оцінювання якості мовних сигналів, спотворених забарвленими шумами. Проте одержані таки чином результати також не є остаточними, оскільки якість та розбірливість мовних сигналів є поняттями хоча і тісно пов'язаними, але не тотожними.

Для усунення зазначених недоліків в Лабораторії акустичної експертизи та корекції кафедри акустики та акустoeлектроніки НТУУ «КПІ ім. І.Сікорського» реалізовано навчально-дослідницький проект [7], в рамках якого запропоновано дев'ять артикуляційних таблиць звукосполучень української мови та створено комплекс комп'ютерних програм для автоматизації артикуляційних випробувань [8], [9].

Мета даної роботи полягає у висвітленні особливостей цього проекту, які не знайшли достатнього відображення в працях [8], [9], а також у поглибленому аналізі одержаних результатів.

Організація випробувань

Можливість подвійного сприйняття окремих фонем було передбачено в оригінальних таблицях стандарту 50840-95, призначених як для неавтоматизованого (рис. 2а), так і автоматизованого (рис. 2б) варіантів артикуляційних випробувань.

ЗАЛЬ	ФИЦ	Д	Б
Ы	В	ГЕТЬ	РАПЬ
ШИНЬ	ЧТАФ	ЖЬ	ПРЁТ
		БУШ	Ц

а

__ЯК/__АК
__ЮЛЬ/__УЛЬ
__ЮЩ/__УЩ
__ЯТ/__АТ
__ВУМ

б

Рис. 2. Фрагменти таблиць стандарту 50840-95, призначені для неавтоматизованого (а) та автоматизованого (б) варіантів артикуляційних випробувань

Шляхом адаптації артикуляційних таблиць для неавтоматизованого варіанту артикуляційних випробувань до української мови, в межах проекту [7] було розроблено дев'ять прототипів артикуляційних таблиць. При цьому кожна таблицю представлено двома варіантами (рис. 3) для врахування можливості подвійного сприйняття окремих фонем.

Термін «прототип» в даному випадку використано, щоб підкреслити попередній характер сформованих таблиць, при створенні яких не в повній мірі враховано фонетичні особливості української мови. Так, зокрема, не враховано

відмінність частотних характеристик фонем української та російської мов, а також притаманність лише українській мові фонем у вигляді африкат «дж» та «дз».

заль	фіц	геть	рап
шин	тав	бущ	прьот

а

заль	фіц	хеть	раб
шин	таф	бущ	прьод

б

Рис. 3. Фрагменти двох варіантів розроблених артикуляційних таблиць (а, б)

При експериментальних випробуваннях цих таблиць та комплексу комп'ютерних програм використовувалися як чисті (неспотворені) мовленнєві сигнали, так і сигнали, спотворені шумом та реверберацією.

Записи чистих сигналів робилися в заглушеному приміщенні кафедри акустики та акустoeлектроніки НТУУ «КПІ ім. І. Сікорського». Для записів використовувалися зовнішня звукова карта PRESONUS AudioBox USB, мікрофон Superlux ECM 999 та програмний звуковий редактор Audacity. Запис виконувався із частотою дискретизації 22050 Гц та бітовою глибиною 16 біт. Звукосполучення зачитувалися із застосуванням словесного оточення, наприклад, звукосполучення «няк» зачитувалося як «Запишіть няк тепер».

Прослуховування відбувалося для чотирьох ситуацій:

- 1) чиста мова;
- 2) мова, спотворена шумом;
- 3) мова, спотворена реверберацією;
- 4) мова, спотворена шумом та реверберацією.

В першому випадку слухачами почергово прослуховувалися записи звукосполучень 3-х артикуляційних таблиць, кожна з яких містила 50 звукосполучень.

В другому випадку прослуховувалися звукосполучення, спотворені адитивним шумом із відношеннями сигнал-шум -10, 0 та +10 дБ. При цьому використовувалися моделі білого, рожевого та коричневого шумів, маскувальна властивість яких вважається досить добре вивченою [1].

В третьому випадку моделювалася, шляхом згортки сигналів чистої мови із імпульсними характеристиками різних приміщень, реверберована мова для часів реверберації від 0,3 до 2,7 с, а в четвертому випадку розглядався випадок спільної дії рожевого шуму та реверберації.

Моделювання сигналів, спотворених адитивним шумом та реверберацією виконувалося за загальним алгоритмом:

$$y(t) = x(t) \otimes h(t) + n(t), \quad (1)$$

де $x(t)$ є мовний сигнал, $h(t)$ є імпульсна характеристика приміщення, $n(t)$ є адитивний шум, \otimes є символ згортки. В окремому випадку дії виключно шумової завади алгоритм (1) спрощується:

$$y(t) = x(t) + k \cdot n(t), \quad (2)$$

де $k = 10^{0.05(SNR-SNR_0)}$, $SNR = 10 \lg(D_x/D_n)$ є «початкове» значення відношення сигнал-шум для чистої мови $x(t)$ та синтезованого «початкового» шуму $n(t)$, D_x та D_n є дисперсіями сигналу $x(t)$ та шуму $n(t)$, відповідно, SNR_0 є бажане SNR для суміші (2).

В окремому випадку сумісної дії шуму та реверберації, алгоритм (1) дещо ускладнюється з метою врахування особливостей структури імпульсної характеристики приміщення:

$$\begin{aligned} h(t) &= h_e(t) + h_l(t) \\ h_e(t) &= \begin{cases} h(t), & t \in 0..50 \text{ ms}; \\ 0, & t \notin 0..50 \text{ ms}, \end{cases} \\ h_l(t) &= \begin{cases} h(t), & t > 50 \text{ ms}; \\ 0, & t \leq 50 \text{ ms}, \end{cases} \end{aligned}$$

де $h_e(t)$ описує ранні відбиття, а $h_l(t)$ описує пізню реверберацію. Тоді «початкове» $SNR = 10 \lg D_{x_e}/D_n$, де D_{x_e} є дисперсія сигналу $x_e(t) = x(t) \otimes h_e(t)$.

Прослуховування виконувалося через навушники 26 студентами середнім віком 22 роки. Три студенти із зазначеного загалу виконали роботу із відхиленнями від завдання, а саме, прослуховуючи сигнали через акустичні монітори (або через комп'ютерні колонки, які в подальшому також будемо називати акустичними моніторами). Тому одержані результати було вирішено розділити на 2 групи: прослуховування через навушники та прослуховування через акустичні монітори.

В процесі артикуляційних випробувань слухачі мали зафіксувати, за допомогою комп'ютерної клавіатури, почуте звукосполучення. Даний спосіб фіксації почутого принципово відрізняється від такого, що пропонується в стандарті 50840-95. Це зроблено свідомо з метою максимального наближення процедури й результатів артикуляційних випробувань до неавтоматизованого варіанту випробувань, який доречно вважати еталонним з огляду на його первинний характер.

Обробка одержаних результатів полягала в оцінюванні розбірливості мови шляхом підрахунку частки правильно розпізнаних звукосполучень. При цьому правильно розпізнаним вважалося таке зафіксоване звукосполучення, символічний склад якого повністю співпадав із таким хоча б для одного із двох текстових варіантів артикуляційних таблиць.

Результати випробувань та їх обговорення

На рис. 2-4 представлено графіки залежності усереднених оцінок розбірливості мови, спотвореної шумом та реверберацією, для випадку прослуховування через навушники.

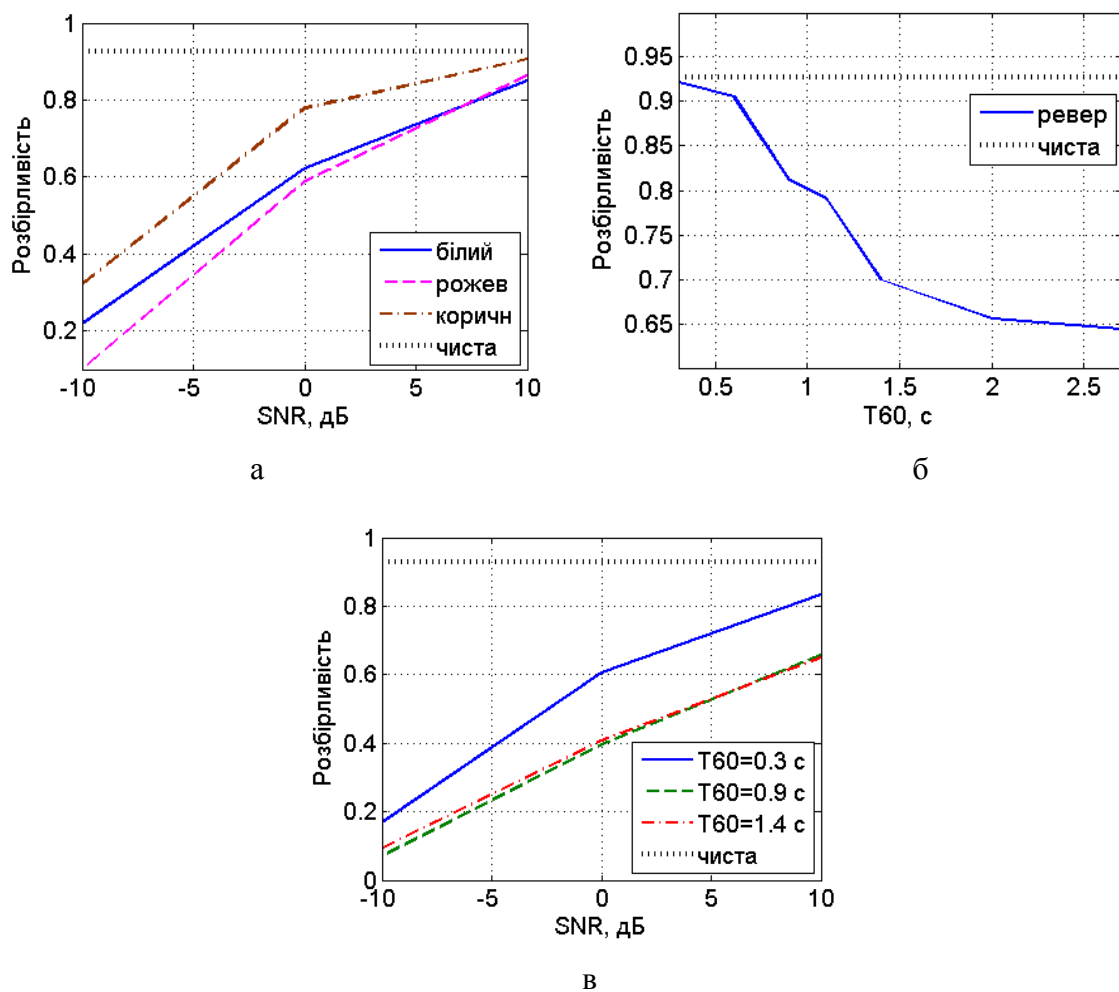
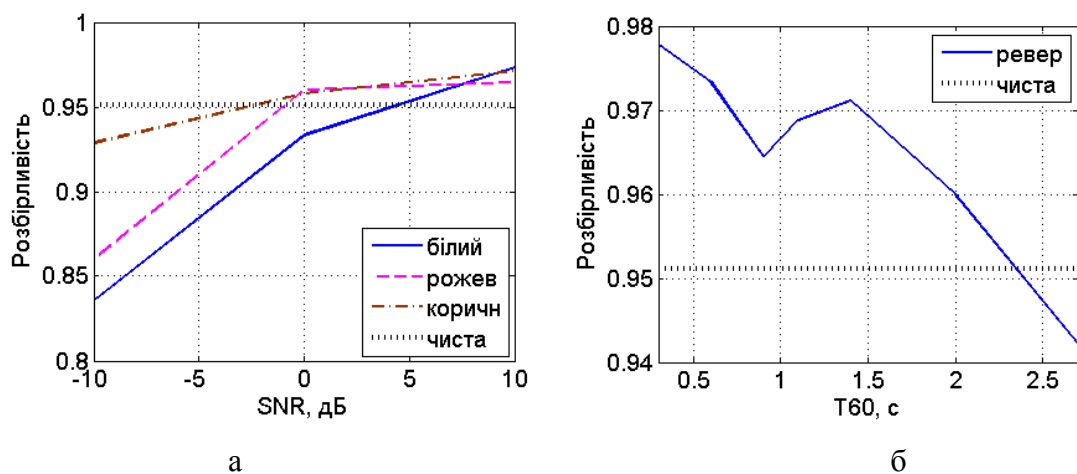


Рис. 4. Оцінки розбірливості при слуханні через навушники: шум (а), реверберація (б), рожевий шум+реверберація (в)

Аналогічні результати для випадку прослуховування через акустичні монітори представлено на рис. 5.



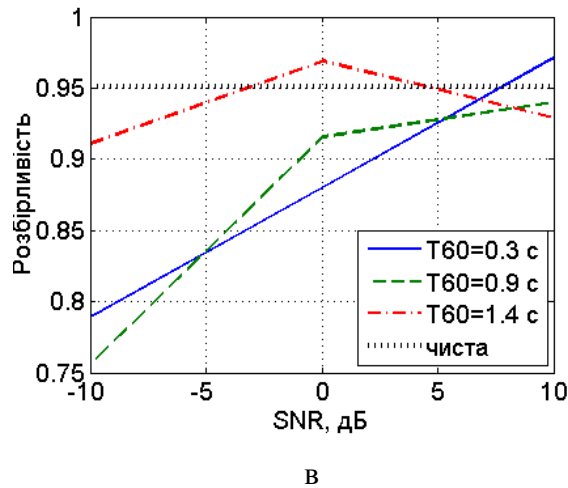


Рис. 5. Оцінки розбірливості при слуханні через акустичну систему: шум (а), реверберація (б), рожевий шум+реверберація (в)

Аналіз графіків рис. 4 та 5 свідчить про те, що маскувальні властивості шуму та реверберації залежать не тільки від відношення сигнал-шум, забарвленості шуму та часу реверберації, але й можуть суттєво залежати від способу прослуховування.

Так, у випадку шумової завади, при переході від прослуховування на навушники до прослуховування через акустичні монітори, розбірливість піднялася із 0,1-0,3 до 0,85-0,93 при відношенні сигнал-шум мінус 10 дБ (стандартне відхилення оцінок розбірливості було близьким в середньому до 5%). Це означає, що відношення сигнал-шум піднялося на 17-20 дБ при прослуховуванні сигналів через акустичні системи, порівняно із прослуховуванням через навушники.

Одержаний приріст відношення сигнал-шум дещо дивує, оскільки є надто вже великим, хоча сам факт зростання розбірливості є очікуваним. Дійсно, при прослуховуванні зашумленого звуку, записаного в режимі «моно», через акустичну систему, що складається із 2-х випромінювачів звуку, на вході кожного вуха сигнали прямого звуку від кожного випромінювача є когерентними, тоді як шуми можна вважати некогерентними. В результаті відношення сигнал-шум (для дії прямого звуку) на вході кожного вуха має збільшитися на 3 дБ порівняно із випадком використання навушників. Значну

роль відіграють ранні відбиття звуку в приміщеннях, аналіз ролі яких [11], [12], [13] свідчить, що за рахунок дії ранніх відбиттів ефективно відношення сигнал-шум може підвищуватися ще на 6-9 дБ. Нарешті, виграш ще приблизно у 3 дБ можна пояснити бінауральним прослуховуванням [14]. Просте підсумовування дії зазначених чинників зрештою дає нам виграш у 12-15 дБ. Це вже досить близько до одержаного нами виграшу 17-20 дБ, хоча не дістає ще приблизно 3-5 дБ.

Найбільш імовірною причиною цієї, досить значної, різниці видається недосконалість організації експериментальних досліджень, а саме:

- неповне врахування особливостей української мови при створенні артикуляційних таблиць;
- обмеженість вибірки дикторів;
- недостатньо висока якість звукового варіанту артикуляційних таблиць;
- недостатня кількість текстових варіантів артикуляційних таблиць;
- недоліки інтерфейсу програмного забезпечення.

Виходячи з цього, доцільно в майбутньому усунути зазначені недоліки й перевірити справедливність припущення, що прослуховування зашумлених сигналів через пару акустичних моніторів у звичайних приміщеннях не може призводити до виграшу у відношенні сигнал-шум, вищого за 12-15 дБ.

Слід припустити, що ще одним важливим чинником може бути психофізичний стан слухачів. Підґрунтям для такого припущення є той факт, що при пізніших повтореннях експерименту із прослуховування сигналів через акустичні монітори в одній із кімнат, де було одержано аномально високу розбірливість мови, нові оцінки розбірливості не перевищували 0,3-0,4 для відношення сигнал-шум мінус 10 дБ. Обговорення із слухачем можливих причин цього факту показало, що перше прослуховування, де було одержано аномально високу оцінку розбірливості, відбувалося в стані певного емоційного підйому. Так чи інакше, але гіпотеза щодо зазначеного чинника має право на існування.

Нарешті, в подальшому доцільно детально проаналізувати акустичні характеристики кімнат, в яких одержано зазначений вище аномально великий приріст розбірливості мови, з метою виявлення можливих особливостей геометрії цих кімнат.

Повертаючись до ситуації прослуховування сигналів через навушники, зазначимо, що графіки рис. 4а досить добре узгоджуються із відомим результатами [1], [2], [7], [15] в області середніх та високих значень SNR (0-10 дБ). Проте в області малих значень SNR (< -5 дБ) маскувальна здатність білого шуму виявилася кращою за таку для коричневого шуму. Це суперечить попереднім прогнозним оцінкам [1], [2], тому доречно також в подальшому дослідити причини зазначеної невідповідності.

Що стосується дії ревербераційної завади, підтверджено очікування щодо зниження розбірливості мови із ростом часу реверберації. Разом із тим, бачимо, що прослуховування через монітори також значно підвищило розбірливість мови. Так, розбірливість зростає із 65% до 94% для часу реверберації $T_{60}=2,7$ с, а для часу реверберації $T_{60}=0,3-2$ с розбірливість навіть перевищила таку для чистої мови на 1-3%. Зазначимо, що в даному випадку стандартне відхилення оцінок розбірливості було досить великим й близьким до 10%, що пояснюється малою кількістю слухачів.

Аналіз оцінок розбірливості для випадку спільної дії шуму та реверберації також свідчить про суттєве зростання розбірливості мови при прослуховуванні її через акустичні монітори (стандартне відхилення оцінок розбірливості не перевищувало 10%).

Підсумовуючи, зазначимо, що за результатами експериментальних досліджень розроблену автоматизовану систему артикуляційних випробувань можна визнати працездатною. Що стосується її якості, перевірка виявила ряд певних недоліків й допомогла сформулювати відповідні рекомендації із їх усунення:

- при створенні артикуляційних таблиць слід більше уваги приділяти фонетичним особливостям української мови;

- перед записом звукових варіантів артикуляційних таблиць необхідно не тільки відповідним чином налаштовувати апаратно-програмне забезпечення, але й ретельно інструктувати та тренувати дикторів, застерігаючи їх від підкреслення звукосполучень за допомогою пауз або шляхом збільшення гучності;
- замість двох текстових варіантів артикуляційних таблиць доцільно використовувати три варіанти, що полегшує врахування не тільки неоднозначності слухового сприйняття окремих фонем, але й неоднозначності введення із клавіатури почутих звукосполучень;
- прослуховування звукосполучень має виконуватися у випадковому порядку, що зменшить ризик запам'ятовування звукосполучень;
- необхідною є можливість корекції введених із клавіатури звукосполучень, оскільки стомлення слухача призводить до збільшення частоти помилкових натиснень клавіатури;
- доцільним є обмеження обсягу візуальної інформації, яка надається слухачу на моніторі після прослуховування кожної артикуляційної таблиці, що дозволяє зменшити ризик одержання поліпшених результатів за рахунок якихось аналітичних міркувань.

Висновки:

Запропоновано та експериментально перевірено автоматизовану систему артикуляційних випробувань. Результати експериментальних досліджень свідчать, з однієї сторони, про працездатність та досить високу якість розробленої системи, що робить можливим її використання як в учбових цілях, так і при розробці відповідного вітчизняного стандарту. З іншої сторони, в процесі експериментальної перевірки виявлено феномен аномально високої розбірливості мови при прослуховуванні спотворених сигналів через акустичні системи. Крім того, при малих відношеннях сигнал-шум виявлено невідповідність одержаних функціональних залежностей попереднім прогностичним оцінкам.

Вироблено рекомендації стосовно подальшої модернізації розробленої системи артикуляційних випробувань, спрямованої на підвищення достовірності артикуляційних випробувань.

Література до підрозділу 1.7.1

1. Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации. Монография. – Киев, "Имекс-ЛТД", 2008. – 420 с.
2. Продеус А., Дидковский В., Дидковская М. Акустическая экспертиза и коррекция коммуникационных каналов. Монография. – LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, Saarbrücken, Deutschland, 2017, ISBN: 978-3-330-04591-0
3. Архипова О., Журавльов В., Кумейко В. Артикуляційні таблиці слів української мови. - Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, вип 2(19) ., 2009. - С. 13-17.
4. Архипов А.Е, Архипова Е.А. Анализ и обработка данных артикуляционных испытаний. - Научно-практический журнал «Защита информации» № 4, 2012. - С. 34-42.
5. Михайлов В.Г. Диагностические артикуляционные таблицы. - Акустический журнал, 2002, том 48, №5. - С. 705-712.
6. ANG-2200. Acoustic Noise Generator. Owner's Guide. Research Electronics, Intl. 455 Security Place Algood, TN 38506, USA, 33 p. [Online]. Available: https://reiusa.net/wp-content/uploads/2017/11/ANG_Manual_revG.pdf
7. Продеус А. М., Вітик А. В., Діденко Д. Ю. Суб'єктивне оцінювання якості та розбірливості мовленнєвих сигналів, спотворених синтезованими шумами [Subjective evaluation of quality and intelligibility of speech distorted by synthesized noise]. Microsystems, Electronics and Acoustics, 2017, vol. 22, no. 6, p. 56-63. ISSN 2523-4447.

8. Лабораторія акустичної експертизи та корекції. Проект: "Автоматизація суб'єктивного оцінювання розбірливості мовлення на тлі шумової та ревербераційної завад". Доступ: <http://acoustic.kpi.ua/index.php?id=2206>

9. A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi, I. Shherbenko. Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility (в печати).

10. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi. Automated Subjective Evaluation of Speech Intelligibility in Noise and Reverberation (в печати).

11. J. Bradley, H. Sato, M. Picard, "On the importance of early reflections for speech in rooms," Journal of Acoustical Society of America, Vol.113, No.6, June 2003, pp. 3233-3244.

12. J. S. Bradley, "Speech intelligibility studies in classrooms," Journal of the Acoustical Society of America, Vol. 80, No. 3, 1986-09, pp. 846-54.

13. W. Yang and J. S. Bradley, "Effects of room acoustics on the intelligibility of speech in classrooms for young children," J. Acoust. Soc. Am. 125 (2), February 2009, pp.922-933.

14. I. Arweller, J. Buchholz, T. Dau, Speech intelligibility enhancement by early reflections. - Proceedings of 2nd International Symposium on Auditory and Audiological Research (ISAAR 2009), August 2009, Elsinore, Denmark.

15. Макаров Ю. К., Хорев А. А. К оценке эффективности защиты акустической (речевой) информации. Информационная безопасность № 4, 2005.

1.7.2. Об'єктивне оцінювання розбірливості мови

Після обчислення ІХ приміщення, стає можливим оцінювання розбірливості мови модуляційним [16], [17] або формантно-модуляційним [18], [19] методом. Першим кроком такого оцінювання є обчислення, за формулою Шредера [20], так званих коефіцієнтів передачі модуляції:

$$m_{ki} = \frac{\left| \int_0^{\infty} h_k^2(t) \exp(-j2\pi F_i t) dt \right|}{\int_0^{\infty} h_{ki}^2(t) dt}, \quad (3)$$

де $h_k(t)$ - результат фільтрації функції $h_r(t)$ смуговим фільтром (на практиці використовують 7 октавних фільтрів із центральними частотами від 125 Гц до 8 кГц); F_i - частота модуляції (на практиці використовують 14 значень F_i , в діапазоні від 0,63 Гц до 12,5 Гц).

Наступними кроками є обчислення ефективних відношень сигнал-шум для кожної смуги та кожної модуляційної частоти:

$$E_{ki} = 10 \lg \frac{m_{ki}}{1 - m_{ki}}, \quad (4)$$

середнього індексу передачі модуляції

$$MT_k = \frac{1}{14} \sum_{i=1}^{14} T_{ki}, \quad (5)$$

$$T_{ki} = \begin{cases} 0, & E_{ki} < -15; \\ (E_{ki} + 15)/30, & -15 \leq E_{ki} \leq +15; \\ 1, & E_{ki} > +15; \end{cases}$$

Останнім кроком, згідно модуляційного методу [17], є обчислення індексу передачі мови:

$$STI = \sum_{k=1}^7 \alpha_k \cdot MT_k - \sum_{k=1}^6 \beta_k \cdot \sqrt{MT_k \cdot MT_{k+1}}, \quad (6)$$

де α_k - вагові коефіцієнти, β_k - коефіцієнти надмірності, значення яких для октавних смуг із центральними частотами f_0 наведено в табл. 1.

Таблиця 1 – Вагові коефіцієнти та коефіцієнти надмірності для октавних смуг

$f_{0k}, \text{Гц}$	125	250	500	1000	2000	4000	8000
α_k	0,085	0,127	0,230	0,233	0,309	0,224	0,173
β_k	0,085	0,078	0,065	0,011	0,047	0,095	-

Згідно формантно-модуляційного методу [18], на останньому кроці обчислюють словесну розбірливість:

$$W = \begin{cases} 1,54 \cdot A^{0,25} [1 - \exp(-11 \cdot A)], & A < 0,15; \\ 1 - \exp\left(\frac{11 \cdot A}{1 + 0,7 \cdot A}\right), & A \geq 0,15, \end{cases} \quad (7)$$

$$A = \sum_{k=1}^7 p_k \cdot P_k(E_k), \quad (8)$$

де A - формантна розбірливість; p_k - імовірність знаходження формант в k -й смузі частот; $P_k(E_k)$ - коефіцієнт сприйняття мови. В працях Н.Б. Покровського [22] та М.А. Сапожкова [23] коефіцієнти сприйняття відрізняються та вважаються незалежними від смуги частот: $P_k(E_k) = P(E_k)$. В [24] викладено результати досліджень причин такої ситуації та показано доцільність використання уточнених коефіцієнтів сприйняття, форма яких залежить від смуги частот.

Метою даної роботи є оцінювання індексу STI та словесної розбірливості W для різних точок кількох лекційних приміщень та зіставлення отриманих результатів між собою, а також із результатами суб'єктивного оцінювання розбірливості мови. Такі порівняння дозволять зробити висновок щодо достовірності отриманих результатів та спроможності розробленого комплексу «Штучна голова» [1] до об'єктивного вимірювання розбірливості мови.

Організація досліджень

На відміну від німецького вимірювального комплексу, де використовувалася професійна аудіоапаратура [5], розроблений комплекс «Штучна голова» [1] містить електроакустичну апаратуру різної якості, а саме: побутову активну колонку Genius SP-HF 2.0 500 (14 Вт, 65 Гц - 20 кГц, сигнал-шум 73 дБ), всеспрямовані вимірювальні конденсаторні мікрофони Superlux ЕСМ-999 (20 Гц - 20 кГц, динамічний діапазон 106 дБ, сигнал-шум 70 дБ), зовнішню звукову карту Steinberg UR242 (24 біт, 192 кГц).

Запис сигналів виконувався із частотою дискретизації 44,1 кГц та глибиною квантування 24 біти. Два мікрофони розташовувалися в зонах вушних раковин штучної голови на відстані 1 см від місцезнаходження вушного каналу.

Основою тестового сигналу слугувала mls-послідовність [25] із ступенем 16 (тобто mls-послідовність містила 2^{16} вибірок), що відповідає довжині сигналу 1,49 с при частоті дискретизації 44,1 кГц. Ця mls-послідовність при випроміненні повторювалася 17 разів, що дозволяло при обчисленні ІХ приміщення усереднювати 16 останніх сплесків оцінки ВКФ для підвищення відношення сигнал-шум на 12 дБ.

Для вимірювань було обрано 2 аудиторії, малого та середнього розміру (рис. 2).

Характеристики малої аудиторії №438: розміри 7х7х3,1 м, два вікна, двері, задня стінка приміщення зашклена, три ряди із 6 парт в кожному ряді та стіл викладача, наповненість приміщення під час запису сигналів - 11 студентів та викладач. Дистанції від гучномовця до номерів 1, 2 та 3 точок, де розміщувалася ШГ, 2 м, 4 м та 6 м, відповідно.

Характеристики середньої аудиторії №209: розміри 10х15х3,1 м, чотири вікна, двері, дві книжних шафи, шафа-гардероб, три ряди із 9 парт в кожному ряді, стіл викладача та 2 додаткових вільних столи, наповненість приміщення під час запису сигналів - 9 студентів та викладач. Дистанції від гучномовця до номерів 1, 2, 3, 4, 5 та 6 точок, де розміщувалася ШГ, 3 м, 6 м, 9 м, 14 м, 9 м та 9 м, відповідно.

Результати досліджень

На рис. 3 наведено графіки ІХ розглянутих приміщень, із результатами вимірювання часу реверберації. Графіки наведено лише для лівого каналу, оскільки для правого каналу вони є подібними. Обвідну ІХ, побудовану за методом Шредера [26], використано для обчислення часу реверберації, оцінки якого представлені на рис. 3.

Порівнюючи ІХ аудиторій 438 (рис. 3б) та 209 (рис. 3в), бачимо, що в аудиторії 209 спостерігається біля 3-х потужних ранніх відбиттів на інтервалі часу до 30 мс, та ще один виразний сплеск в момент часу 0,27 с, тоді як спад енергії в аудиторії 438 є більш плавним. Час реверберації в обох приміщеннях виявився досить великим та близьким до 1 с, що у випадку аудиторії 209 пояснюється слабким заповненням порівняно великого приміщення, а у випадку аудиторії 438 пояснюється наявністю скляної задньої стінки.

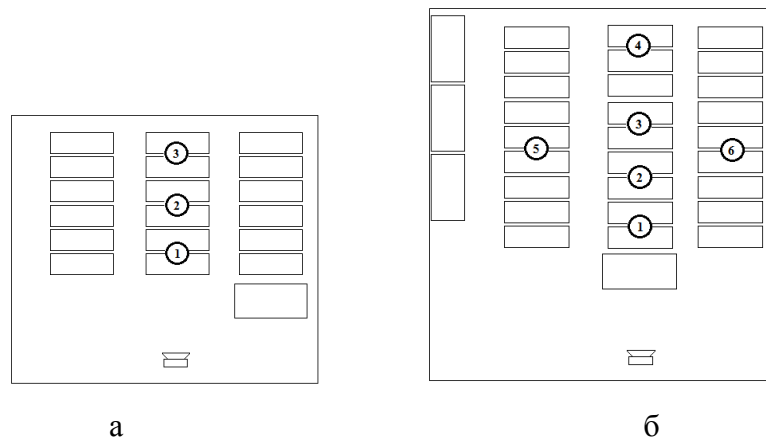


Рис. 2. Плани малої аудиторії 438 (а) та середньої аудиторії 209 (б)

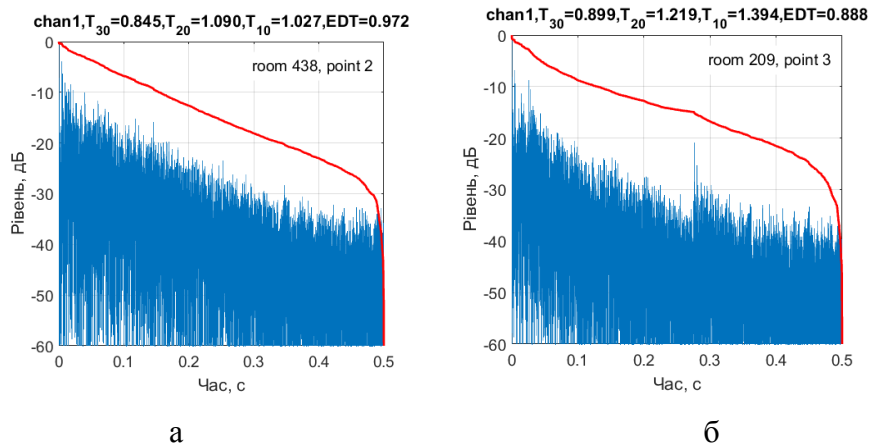


Рис. 3. ІХ та час реверберації для аудиторій 438 (а) та 209 (б)

Результати оцінювання характеристик $C_{50} = 10\lg(E_0^{50}/E_0^{\infty})$ та $ERB = 10\lg(E_0^{50}/E_0^{10})$ наведено на рис. 4 та 5, відповідно, де $E_a^b = \int_a^b h^2(t)dt$, a та b - час в мілісекундах. Оскільки для $t > 0,5$ с рівень шумової завади переважає над рівнем реверберації (рис. 3), при обчисленнях C_{50} в якості ∞ використано значення 0,5 с.

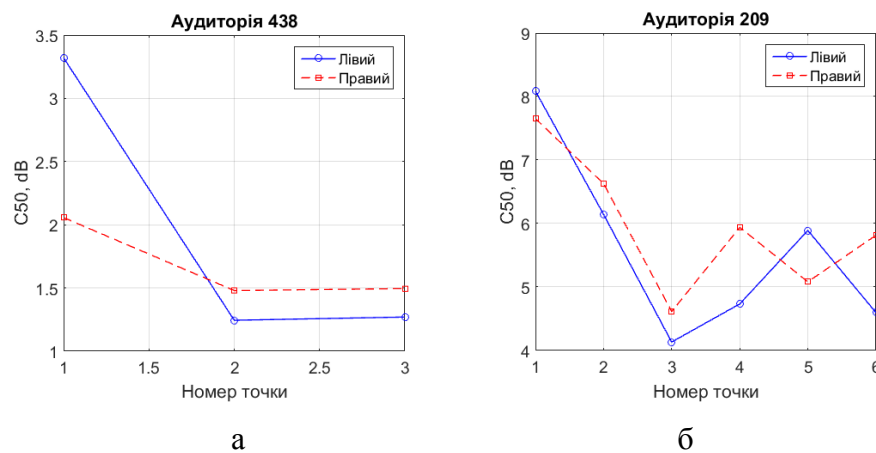


Рис. 4. Оцінки параметру C_{50} для аудиторій 438 (а) та 209 (б)

Згідно отриманим оцінкам C_{50} , чіткість мовлення в аудиторії 209 має бути помітно вищою, ніж в аудиторії 438, причому ця чіткість має швидко спадати із наближенням до середини кімнати, після чого має стабілізуватися на певному рівні. Цікаво, що, згідно оцінкам C_{50} , в аудиторії 209 чіткість мови біля задньої стінки (точка 4) є помітно вищою, ніж в середині кімнати (точка 3). В аудиторії 438 подібний ефект також спостерігається, але є ледь помітним.

Поведінка показника ERB , що характеризує співвідношення енергій ранніх відбиттів та прямого звуку, також підтверджує існування зазначеного ефекту. Дійсно, якщо вважати, що прямий звук є кращим з точки зору розбірливості мови, порівняно із ранніми відбиттями, тоді, згідно рис. 5а, бачимо, що в аудиторії 438 розбірливість в лівому каналі є вищою біля задньої стінки (точка 3), порівняно із серединою приміщення (точка 2). Проте результати оцінювання ERB в правому каналі свідчать про протилежне. В аудиторії 209 (рис. 5б)

ситуація є не такою суперечливою: в обох каналах, розбірливість мови біля задньої стінки (точка 4) мала би покращитися, порівняно із серединою приміщення (точка 3).

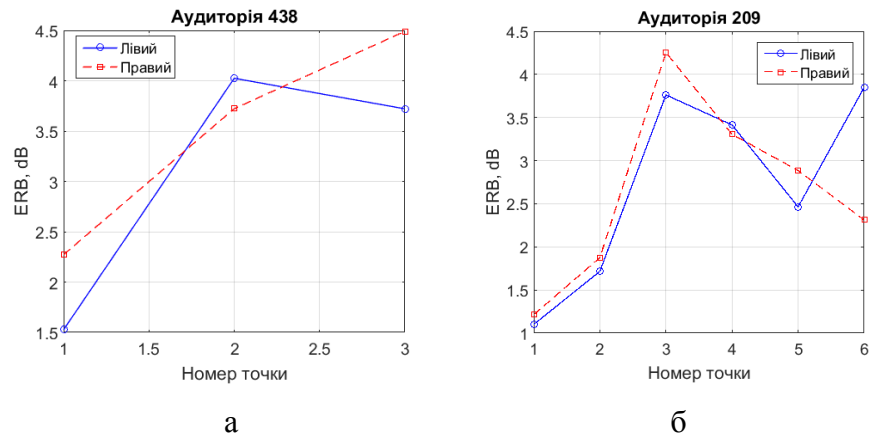


Рис. 5. Оцінки параметру ERB для аудиторій 438 (а) та 209 (б)

Зважаючи на певну неоднозначність отриманих оцінок C50 та ERB, які є непрямыми мірами розбірливості мови, цікаво оцінити розбірливість мови у вказаних аудиторіях модуляційним та формантно-модуляційним методами із використанням (3)-(7).

Результати оцінювання розбірливості мови модуляційним методом представлено на рис. 6, а формантно-модуляційним методом – на рис. 7, при цьому оцінки, отримані формантно-модуляційним методом, були розраховані із використанням уточнених коефіцієнтів сприйняття, форма яких залежить від смуги частот [24].

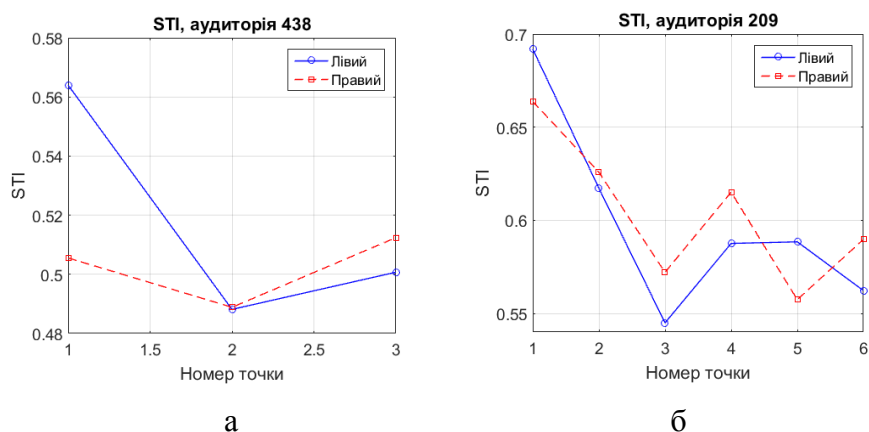


Рис. 6. Оцінки індексу STI для аудиторій 438 (а) та 209 (б)

Неважко бачити хороше узгодження оцінок розбірливості мови, отриманих модуляційним та формантно-модуляційним методами, а також узгодження їх із оцінками параметру C50 (рис. 4). Слід, однак, зазначити, що ефект підвищення розбірливості мови біля задньої стінки аудиторії 438, порівняно із серединою приміщення, є значно виразнішим для оцінок розбірливості мови (рис. 6а та 7а), ніж для коефіцієнту C50 (рис. 4а). Що стосується аудиторії 209, де було виконано вимірювання ще й біля бокових стінок (точки з номерами 5 та 6), бачимо, що біля бокових стінок аудиторії розбірливість мови є також вищою, ніж в центрі приміщення. Зазначимо, що ефект підвищення розбірливості мови біля стінок приміщення було виявлено також в [14] шляхом суб'єктивного оцінювання розбірливості мови.

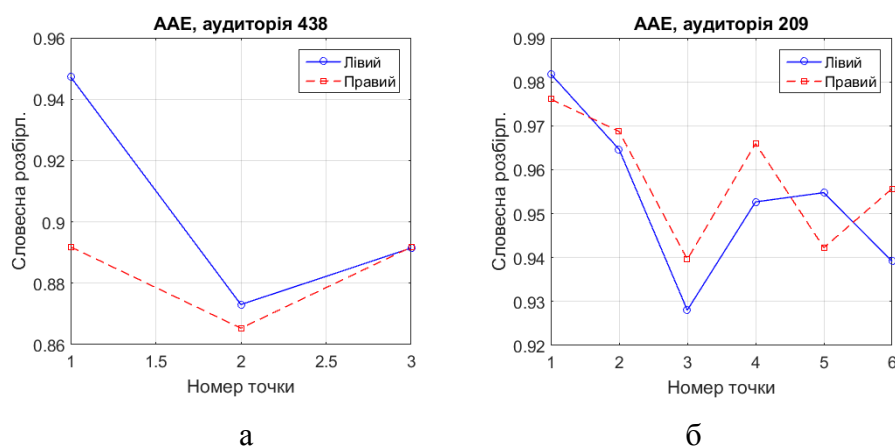


Рис. 7. Оцінки словесної розбірливості для аудиторій 438 (а) та 209 (б)

Даному ефекту можна дати досить просте фізичне пояснення. Дійсно, до вух людини, що сидить біля стінки, майже одночасно (в наших випадках різниця в часі становила 3-5 мс) приходять прямий та відбитий від стінки звуки. При цьому потужність відбитого від стінки звуку є ненабагато меншою, порівняно із потужністю прямого сигналу. На низьких частотах (100-200 Гц) сигнали є практично когерентними, тому при близьких рівнях прямого та відбитого звуку на цих частотах можна очікувати підвищення рівня результуючого звуку до 6 дБ, а на високих частотах, де сигнали є практично некогерентними, рівень звуку може підвищитися до 3 дБ. Зауважимо, що хоча ці оцінки є дуже приблизними,

проте вони добре узгоджуються із результатами, наведеними в [11], [12]. Що стосується середини приміщення, відбитий від стінки звук приходить із більшою затримкою відносно прямого звуку (в наших випадках 15-20 мс) та є послабленим за рахунок розповсюдження в просторі. Оскільки про когерентність прямого та відбитого звуків в даному випадку вже не йдеться, то очевидно, що рівень сумарного сигналу має бути меншим, ніж біля стінки. Крім того, слід вказати ще й на те, що через помітну затримку між прямим та відбитим сигналом кімната фактично перетворюється в гребінчастий фільтр (комб-фільтр) [27] із суттєво нерівномірною амплітудно-частотною характеристикою. Таким чином, сумарний звук є ще й спотвореним, що також не може сприяти хорошій розбірливості мови в середині кімнати.

Узагальнюючи, можна сказати, що в аудиторіях, подібних розглянутим в даній роботі, користь ранніх відбиттів, за критерієм розбірливості мови, є різною в різних частинах приміщення, й біля стін приміщення вона є більшою, ніж в середині приміщення.

Оскільки в даній роботі, за браком місця, не приділено уваги аналізу співвідношень розбірливості мови в лівому та правому каналах, в подальшому доцільно ліквідувати зазначений пробіл.

Висновки

Результати оцінювання різних мір розбірливості мови, виконаного із використанням розробленого апаратно-програмного комплексу «Штучна голова», показали, що розбірливість мови в середині аудиторій може бути суттєво меншою за таку біля стінок приміщення. Отримані результати добре узгоджуються із результатами попередніх досліджень, що свідчить про можливість використання розробленого апаратно-програмного комплексу «Штучна голова» для двоканального оцінювання розбірливості мови. В подальшому доцільно приділити увагу аналізу співвідношень розбірливості мови в лівому та правому каналах.

Література до підрозділу 1.7.2.2

1. А. Продеус, М. Дідковська, О. Дворник, Д. Моторнюк. Апаратно-програмний комплекс «Штучна голова». Частина 1. Коригування частотної характеристики тракту. Мікросистеми, електроніка та акустика, 2020 (передано до друку).
2. J. Blauert Ed., The technology of binaural listening. Springer, Berlin–Heidelberg–New York, 2013.
3. HMS II.3 Artificial Head Measurement System with Ear Simulator and Artificial Mouth, Data Sheet, HEAD Acoustics GmbH, 2015.
4. P. Søndergaard and P. Majdak. The auditory modeling toolbox. In J. Blauert, editor, The technology of binaural listening, chapter 2. Springer, Berlin-Heidelberg-New York NY, 2013.
5. M. Jeub, M. Schafer, P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," Proc. of 2009 16th International Conference on Digital Signal Processing, 5-7 July 2009, Santorini-Hellas, Greece. DOI: 10.1109/ICDSP.2009.5201259.
6. Aachen Impulse Response Database. Available on-line: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>
7. Ю.С. Костючок, Л.С. Мартинович, Д.Е. Моторнюк, В.А. Нечитайло, А.В. Храпачевский, А.Н. Продеус. Акустическая паспортизация учебных помещений. - ISSN 1811-4512. Electronics and Communications, 2016, Vol.21, No.2(91). – P.63-70.
8. G. Soulodre, N. Popplewell, and J. Bradley, "Combined effects of early reflections and background noise on speech intelligibility," Journal of Sound and Vibration, vol. 135, No.1, pp. 123-133, 1989.
9. Sound Forge Audio Studio. Available on-line: <https://www.magix.com/ru/muzyka/sound-forge/sound-forge-audio-studio/>
10. E. Habets, N. Gaubitch, and P. Naylor, "Temporal selective dereverberation of noisy speech using one microphone," Proc. of 2008 IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 4577-4580, March-April 2008.

11.J. Bradley, H. Sato, and M. Picard, “On the importance of early reflections for speech in rooms,” J. Acoust. Soc. Am., vol. 113, no. 6, pp. 3233-3244, June 2003. DOI: 10.1121/1.1570439.

12.I. Arweiler, J. Buchholz, and T. Dau, “Speech intelligibility enhancement by early reflections,” Proc. of 2nd Int. Symposium on Auditory and Audiological Research (ISAAR 2009), Elsinore, Denmark, August 2009. ISBN: 87-990013-2-2

13.H. Sato and J. Bradley, “Evaluation of acoustical conditions for speech communication in working elementary school classrooms,” J. Acoust. Soc. Am. 106 (4), Pt. 1, pp. 2064–2077, 2004. DOI: 10.1121/1.2839283.

14.A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, and D. Motorniuk, “Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening,” in “Data-Centric Business and Applications,” Springer, 2019 (in publishing).

15.Тихонов А. Н. О некорректных задачах линейной алгебры и устойчивом методе их решения. ДАН СССР, 1965, т. 163, № 3, с. 591—594.

16.H.J.M. Steeneken, and T. Houtgast, “A physical method for measuring speech-transmission quality,” J.Acoust. Soc. Am. - 1980. - Vol. 67. – P. 318-326.

17.H.J.M. Steeneken, and T. Houtgast, “Validation of the revised STIr method,” Elsevier Speech Communication. – 2002. - Vol. 38. – P. 26-37.

18.А.Н. Продеус, Л.Б. Дронжевская, В.А. Климков, Д.А. Шагитова, «Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 1. Унификация алгоритмов», Электроника и связь, 2010, №6, часть 2. - С.117- 124.

19.А.Н. Продеус, Л.Б. Дронжевская, В.А. Климков, Д.А. Шагитова, «Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 2. Точность и скорость измерений», Электроника и связь. – 2011. – Т. 64. – №6. – С. 16–24.

20.M.R. Schroeder, “Modulation Transfer Functions: Definition and Measurement,” Acustica, Vol. 49, 1981, p. 179-182.

- 21.Дідковський В.С., Дідковська М.В., Продеус А.М. Комп'ютерна обробка акустичних сигналів. Навчальний посібник. – К., «Імекс-ЛТД», 2010. – 420 с.
- 22.Покровский Н.Б. Расчет и измерение разборчивости речи. – М.,Связьиздат,1962,390с.
- 23.Сапожков М.А. Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963. – 472 с.
- 24.Продеус А., Дидковский В., Дидковская М. Акустическая экспертиза и коррекция коммуникационных каналов. Монография. LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, Saarbrücken, Deutschland, 2017, ISBN: 978-3-330-04591-0.
- 25.В.Е. Гантмахер, Н.Е. Быстров, Д.В.Чеботарев. Шумоподобные сигналы. Анализ, синтез, обработка. - СПб.: Наука и Техника, 2005. - 400 с.
- 26.M.R. Schroeder, "New Method of Measuring Reverberation Time," The Journal of the Acoustical Society of America 37, 1187 (1965); <https://doi.org/10.1121/1.1939454>.
- 27.Zolzer U. DAFX: Digital Acoustic Effects. - John Wiley & Sons, 2011. – 614 p.

Лекція 1.8. Акустична експертиза концертних залів

В даному підрозділі представлено матеріалі статті *K. Jacob, Correlation of speech intelligibility tests in reverberant rooms with three predictive algorithms. Proc. of 85th Conception of the Audio Engineering Society, Los Angeles, 1989. - P. 1020-1030.*

Виконано порівняння характеристик результатів трьох об'єктивних тестів на розбірливість мови в 10 залах Бостона різних розмірів, призначених для музичних та мовних презентацій.

У дане дослідження залучені приміщення з високим часом реверберації і представлені складні ситуації посилення звуку. Результати суб'єктивних тестів на розбірливість порівнюються з результатами, отриманими за допомогою трьох опублікованих алгоритмів. Один з алгоритмів, що ґрунтуються на модуляційній передавальній функції, було розглянуто як міжнародний стандарт (IEC Report 268-16, 1988), і його точність підтверджується даними дослідженням. Показано, що другий алгоритм, який базується на відношенні сигнал-шум Лохнера-Бюргера, має таку ж точність при співпадаючих умовах. Показано, що третій алгоритм, є другорядним. Дослідження підтверджує благотворний вплив відбиттів першого порядку на розбірливість мови і негативний ефект, який чинять більш пізні відбиття.

Вступ

В даному дослідженні для кожної комбінації приміщення, випромінювача, позиції слухача були наведені результати суб'єктивних та об'єктивних тестів на розбірливість. Точність кожного з трьох алгоритмів об'єктивного тестування була розглянута саме в цих рамках. Методика, техніка для проведення експериментів, проміжні результати докладно описані.

Три різних використаних джерела перераховані в табл.1.

Було вивчено 10 приміщень, розташованих в центрі Бостона (табл. 2)

Джерела звуку були розташовані в областях приміщення, де зазвичай знаходиться мовець. У приміщеннях, що мають області помосту, на кшталт авансцени, джерела завжди розташовувалися посередині і прямували в бік публіки. У деяких приміщеннях джерела піднімалися приблизно над центром авансцени, за допомогою пневматичної вишки, в інших - містилися на двометрові стійки як показано на рис. 1. У всіх випадках була спроба розмістити джерела в місця, які підходять як для постійних, так і для тимчасових звукових систем.

У кожному приміщенні були обрані дві позиції для слухача. Одна відповідала приблизно одній третині дистанції від джерела до задньої стіни, а друга - в кінці залу для глядачів. Ці позиції були обрані з розрахунком можливого широкого розкиду значень розбірливості. У всіх випадках залежність між джерелами і слухачем була такою, що слухач знаходився всередині сектора з центральним кутом $7,5^0$, рахуючи від головної акустичної осі джерела.

При вивченні ефекту реверберації на розбірливість мови було докладено всіх зусиль для мінімізації впливу інших факторів, що впливають на розбірливість. Джерела були індивідуально відрегульовані, для отримання відгуків рівнозначною частоти (розкид в рамках ± 2 дБ) як вимірювалося на третину-октавних аналізаторі. Смуга частот лежала в межах 100 Гц - 10 кГц. Ледве чутний шум, дзижчання або спотворення джерелами були усунені до проведення тестів.

Фоновий шум був мінімізований як фактор, до значення, що гарантує відношення рівня мовного сигналу до фонового шуму принаймні 30 дБ, встановленим на А-зважуванні. В цілому, потрібно відношення сигнал-фоновий шум принаймні 15 дБ щоб запобігти дії шуму як фактору.

Суб'єктивне тестування

Для встановлення точної оцінки дійсної розбірливості мови, були проведені суб'єктивні тести для кожної комбінації видів приміщення, джерела і

розташування слухача. Суб'єктивні тести були проведені відповідно до ANSI S 3.2-1960 (R1982) [5].

У тесті ANSI, суб'єктам видається комбінація з 20 списків по 50 фонетично збалансованих (ФБ) односкладових слів. При відтворенні слухачеві, слова вставляються в несучу фразу, наприклад: «ви напишіть слово ТЕСТ зараз». Несуча фраза необхідна для симуляції ефекту накладення послідовно вимовлених слів. Рахунок словами ведеться слідуючи більше фонетиці ніж написання. Наприклад «*prayed*» еквівалентно «*raid*». Якщо якась частина слова фонетично сприймається невірно, то і все слово вважається сприйнятим невірно. Відсоток розбірливості визначається як процентне співвідношення фонетично правильно сприйнятих слів і записується як %ФБдійс.

Список слів, відтворений через динаміки, був записаний в заглушеній камері з використанням ненаправленого інструментально проградуйованого мікрофона, встановленого на відстані 0,5 м від диктора, для створення запису мови, що містить дуже малу ревербераційну перешкоду.

Суб'єкти були обрані зі звичайних людей. Пред'являлися вимоги: 3 роки в коледжі і аудіограма в межах норми. При цьому суб'єктам не доводили суть експерименту.

Для перших п'яти приміщень (див. табл. 2) суб'єктивне тестування було проведено безпосередньо в самих приміщеннях. Для другої групи з п'яти кімнат були використані макет голови Sennheiser і комбінація з двох мікрофонів для двох позицій слухача. (Попередньо був проведений експеримент для визначення точності системи бінауральної запису і подальшого відтворення. Результати показують незначні зміни в значеннях розбірливості, якщо тест проведено безпосередньо з суб'єктами в даному приміщенні). Ці бінауральніє записи були програні через головні телефони суб'єктам пізніше. Таким чином тести були бінауральним у всіх випадках.

Тестування було проведено в два підходи, що складаються з п'яти наступних днів кожен. Для кожного з підходів були використані різні суб'єкти. Списки слів були перетасувати щоб уникнути довгострокового навчання

слухачів. Крім того, один і той же список слів ніколи не програвався на тому ж динаміці або для тієї ж позиції слухача день у день. На додаток, два списки ніколи не програвалися послідовно, не дивлячись на інші умови. Сумарна кількість слів відтворених для кожної комбінації приміщення, динаміка і позиції слухача лежить в діапазоні 2000-2800 слів.

Табл. 1

Имя	Тип	Осевая направленность
Soundsphere 2212-1	Ненаправленный излучатель	1.1
Bose 802-II	Восьми-генераторная упорядоченная конструкция	7.3
Electro Voice HR 6040 A (with TL806AX)	Постоянно-направленный рожок	17.7

Табл. 2

Имя	T60, с	Объем, м ³
Berklee Performance Center, Boston	0,9	5450
Coolidge Corouse, Brookline	1	4590
Huntigton theater, Boston	1,1	3190
Saint Bridget's Church, Framingham	2	3810
Nevins Hall, Framingham	3,5	10620
Jordan Hall, Boston	2,2	4530
Mechanics Hall, Worcester	2,2	11582
South End Cathedral, Boston	3,3	59152
The Cyclorama, Boston	3,5	11610
MIT Indoor Track, Cambridge	4,6	42475

Об'єктивні вимірювання

Для кожної комбінації приміщення, виду джерела і позиції слухача, система імпульсних характеристик являє собою передавальні функції від вхідних висновків джерела до слухача, які і записувалися. Для вимірювань, зберігання та аналізу сигналів була використана система отримання даних, що

базується на мікрокомп'ютері. Частота дискретизації була прийнята рівною 9615 Гц, таким чином, представляючи собою практичний смуговий фільтр з центральною частотою приблизно 4 кГц. Для запису кожного імпульсу були використані 8192-вибіркові ділянки буферної пам'яті.

Спрямованість постійно-спрямованого гудка була взята з виробничої відомості. Для ненаправленого і складного випромінювачів, характеристики спрямованості були зняті в заглушеній кімнаті з використанням багато-мікрофонної системи в порядку вимірювання всеспрямованого випромінювання (повна сфера). Спрямованість обох пристроїв не постійна на частотах вище 1000 Гц, завдяки взаємодії декількох вихідних каскадів і дифракційним ефектів. Спрямованість, зазначена в табл.1, являє собою середнє з 1-, 2- і 4-кГц октавних смуг.

Час реверберації було розраховано для кожного приміщення за наступним принципом. Для кожної імпульсної характеристики: 1) сигнал був квадровано (зведено у квадрат), 2) піднесений в квадрат сигнал був інтегрований від максимального часу (понад 800 мс) до нульового часу (так зване зворотне інтегрування), 3) отриманий графік часової залежності звукового тиску конвертувався згідно дБ-ної шкали в форму, якби запис проводився в приміщенні з миттєво вимкнутим постійним шумом, 4) лінійна залежність використовувалася в графіку від 100 до 600 мс після початку загасання, 5) час реверберації розраховувався як час перетину прямої на рівні 60 дБ і спадаючої прямої. Нарешті T60 для декількох ІХ усереднюється. Розкид значень для даної кімнати при цьому становив не більше 3%.

Обсяг приміщення розраховувався з його планів. Відстань від джерела до слухача дорівнював добутку часу затримки від джерела до вимірювального мікрофона на швидкість звуку.

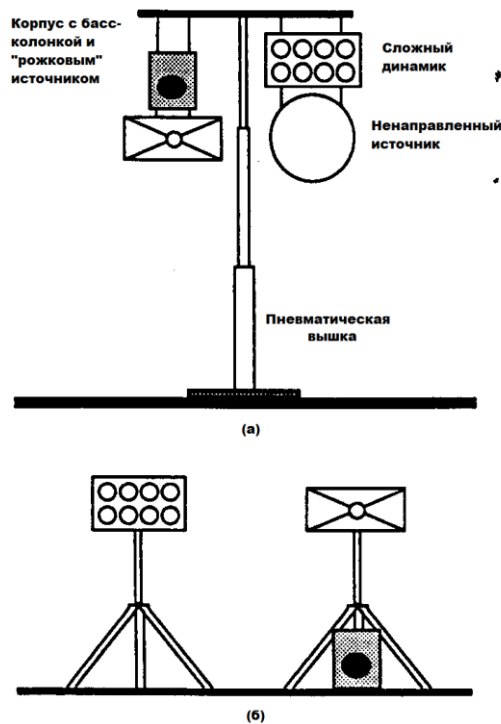


Рис. 1 Две использованных конфигурации расположения динамиков

Алгоритми розрахунку

Модуляційна передатна функція (STI)

Модуляційний метод був використаний як міра того як організм обробляє візуальні сигнали і в цілому для кількісного аналізу характеристик оптичних систем. Модуляційний метод був запропонований Хоутгастом і Стінекеном в 1971 р. як міра розбірливості.

Формула Шройдер дозволяє впровадити модуляційний метод для оцінювання розбірливості мови. Ця формула була реалізована і в даному дослідженні:

$$m(F) = \left| \frac{\int_0^x p^2(t) e^{-j2\pi Ft} dt}{\int_0^x p^2(t) dt} \right| \quad (1)$$

де $m(F)$ – модульіно-передатна функція; $p(t)$ – ІХ системи; F – частота модуляції, Гц.

З виразу 1 видно, що модуляційно-передавальна функція пропорційна величині Фур'є-перетворення квадрата імпульсної характеристики. В цьому експерименті було забезпечено відношення мовної сигнал-фоновий шум більше

30 дБ у всіх випадках. Це дозволяє використовувати вираз 1 з мінімальною помилкою.

Лохнер-Бургерівське відношення сигнал-шум

Лохнер і Бургер представили фундаментальне дослідження про те, як звукова енергія сприймається слуховою системою, особливо в умовах відбиття звуку в приміщенні. Вони з'ясували, відбитий сигнал з рівнем на 5дБ менше ніж у прямого буде повністю інтегруватися вухом, якщо прибуде за час менше або рівне 40 мс (ефект Хааса).

Лохнер і Бургер показали що сумарна енергія ранніх віддзеркалень може сприйматися як частина сигналу, а енергія від пізніх відображень, комбінуючи з фоновим шумом сприймається як шумова складова у відношенні сигнал-шум для розбірливості мови. Інші дослідження мали ту ж основну ідею психо-акустичного характеру відношення сигнал-шум у формуванні об'єктивних мір суб'єктивних вражень [13].

Формула Лохнера-Бургера для відношення сигнал-шум має вигляд:

$$R_e = 10 \log \frac{\int_0^{95 \text{ мс}} p^2(t) \alpha(t, l) dt}{\int_{95 \text{ мс}}^x p^2(t) dt} \quad (2)$$

де R_e – відношення сигнал-шум за Лохнер-Бургером, дБ; $p(t)$ – ІХ системи; $\alpha(t, l)$ – вагова функція, що описує відбиття згідно часовій затримці t та рівню l .

Лохнер і Бургер не привели кінцевих експериментальних даних, що показують точне співвідношення між їх визначенням сигнал-шуму і розбірливості мови. Інші дослідники, серед яких найбільш відзначилися Бредлі [14] і Летхем [15] встановили, що відношення сигнал-шум за Лохнер-Бургером містить всю необхідну інформацію для передбачення розбірливості мови. (У нас це - схожість форми оцінок C50 і STI. - А.П.) Також вони вважали за необхідне, розрахувати свої власні кореляційні функції між об'єктивними і суб'єктивними оцінками.

%ALcons Формула для направленных источников

Пеутц розробив об'єктивний метод визначення розбірливості мови, який підраховує середній відсоток незрозумілих приголосних звуків в списках слів. Результатом є відсоток артикуляційно втрачених приголосних (записується як %ALcons) може бути виражений через %ФБдейст вимірюваний і розраховується з використанням стандарту ANSI.

Пеутц провів експеримент, використовуючи ненаправлене джерело в розмаїтті приміщень. Його дані показали що %ALcons зростає як функція квадрата дистанції джерело-слухач, аж до деякої граничної дистанції, за якої не відбувається подальшого зростання. Дані також показують, що ця гранична дистанція просто співвідноситься з критичної дистанцією для даної комбінації приміщення-джерело, яка визначається як дистанція джерело-слухач на якій пряме поле і стале ревербераційній поле рівні по інтенсивності. Виходячи з цих міркувань, можна записати просту формулу, яка %ALcons:

$$\%AL_{\text{cons pred}} = \begin{cases} \frac{200D^2T^2}{V}, & D < D_L \approx 3,2D_C \\ 9T, & D > D_L \end{cases} \quad (3)$$

де D – відстань джерело-слухач, м; T – час реверберації, с; D_L – гранична дистанція, м; D_C – критична дистанція, м.

Модифікована формула Пеутца, що включає направленість джерела Q :

$$\%AL_{\text{cons pred}} = \begin{cases} \frac{200D^2T^2}{VQ_{AXIAL}}, & D < D_L \approx 3,2D_C \\ 9T, & D > D_L \end{cases} \quad (5)$$

Результати

Результати суб'єктивного тестування

Результати суб'єктивного тестування показані в табл. А 1.

Величини STI були конвертовані в% ФБ, з використанням максимально підходящої в цьому випадку кривої регресії третього порядку, запропонованої Стінекеном і Хоутгастом [9] для широкої різноманітності умов. Формула для даної регресії має вигляд:

$$\%ФБ_{\text{пред}} = -43 + 279,8STI - 31,2STI^2 - 124STI^3 \quad (6)$$

Техніка, використана для конвертації відношення сигнал-шум за Лохнер-Бургером в %ФБ для всіх комбінацій приміщення, джерела і позиції слухача, та ж, що використовувалася Бредлі і Латхамом. Регресійний поліном другого порядку був розрахований, використовуючи порівняння даних %ФБ і відношень сигнал-шум як показано на рис. 2. Рівняння для регресійної кривої може бути використано для конвертації відношень сигнал шум в %ФБ.

Регресійна крива рис.2 представляється виразом:

$$\%ФБ_{\text{пред}} = 79,458 + 2,99R_e - 0,1368R_e^2 \quad (7)$$

Відсоток артикуляційних втрат приголосних звуків %ALcons - це метод суб'єктивного оцінювання мовної артикуляції розбірливості, який відрізняється від методів, які використовують %ФБ. Аналогічно, %ALcons, передбачена формулою, не може на пряму порівнюватися з передбаченою величиною %ФБ. Потрібно співвідношення, щоб конвертувати %ALcons в %ФБ.

Для встановлення співвідношення між %ALcons і %ФБ, списки слів, в даному дослідженні, були повторно озвучені, дотримуючись процедури %ALcons, описаної Пеутцем. Відмінне співвідношення було встановлено між методами, як показано на рис. 3 і використано для конвертації %ALcons в %ФБ.

Регресійна пряма на рис. 3 предствлена виразом:

$$\%ФБ_{\text{дейст}} = 100 - \frac{\%ALcons_{act} + 2,12}{0,7} \quad (8)$$

Для підсумовування, значення STI були конвертовані в %ФБ з використанням записаної раніше формули; відношення сигнал-шум за Лохнер-Бургером були конвертовані за допомогою найбільш підходящого, для зібраних в даному дослідженні даних, графіка; отримані значення %ALcons були конвертовані з використанням співвідношення між %ФБ і %ALcons методами вимірювання розбірливості списків слів. Використовуючи дані переходи, три об'єктивних алгоритми можуть безпосередньо порівнюватися з %ФБ.

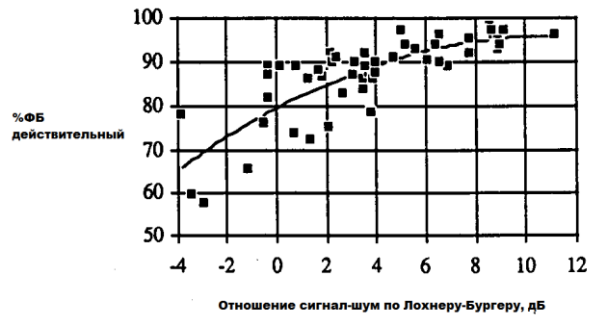


Рис. 2 Отношения сигнал-шум по Лохнеру-Бургеру в сравнении с %ФБдейст и регрессионной прямой

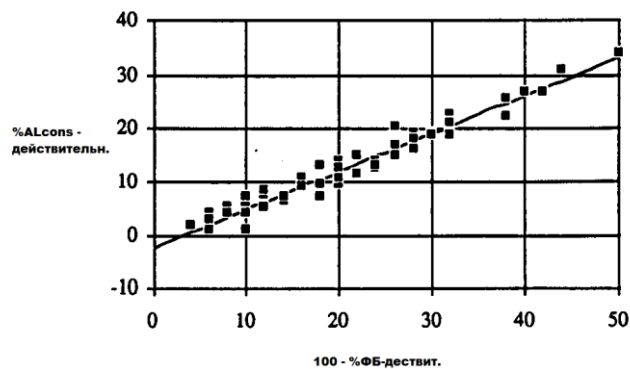


Рис. 3 Диаграмма рассеяния, показывающая соотношение между %ALcons act и %ФБ действ, измеряемых из одних и тех же субъективных тестов.

Залежність об'єктивних та суб'єктивних значень

Діаграми розсіювання %ФБдейст в порівнянні з %ФБпред для кожного з трьох об'єктивних методів показані на рис. 4. Графіки можуть бути інтерпретовані візуально із розумінням того, що в ідеалі, значення, одержувані за допомогою об'єктивного методу, на графіку шикуються в пряму лінію. Точки під прямою лінією представляють об'єктивно отримані значення, які при порівнянні з дійсними значеннями виявилися занадто високими, точки нижче прямої - занадто низькими, відповідно.

Дані показують приблизно однаков розсіювання точок по відношенню до прямої на рис. 4 (а) і (б). Значно більше розсіювання і деякий ухил, можна побачити на рис. 4 (в) для наступних точок над лінією.

Похибки оцінювання

Дані представлені в діаграмах розсіювання, можуть бути вивчені більш ретельно, для встановлення ступеня точності об'єктивно отриманих даних. Найпростіша міра точності - стандартні відхилення, сформовані квадратуванням і

складанням помилок між дійсними і об'єктивно отриманими значеннями і подальшим витяганням кореня квадратного з отриманого результату. Основні стандартні відхилення трьох об'єктивних методів показані на рис. 5. Графік показує, що метод відношень сигнал-шум має найнижче стандартне відхилення, потім дуже близько йде модуляційний метод, тоді як стандартне відхилення для методу % ALcons є значно вищим. Цікаво відзначити, що стандартне відхилення в 7% для модуляційного методу є точно таким же, як наводиться Стінекеном і Хоутгастом. Більш того, також є деякі збіги для стандартного відхилення в 11,5%, знайденого для методу %ALcons, і 17% точністю (10 %ALcons), наведеної Пеутцом [16]. Різні вихідні дані, проте, не роблять ці стандартні відхилення в точності порівнянними.

Середні похибки

Інший спосіб перевірити отримані дані - порахувати середні відмінності між отриманими об'єктивними методами і дійсними значеннями. Такий аналіз дає інформацію, чи містять дані деякий ухил. Наприклад, якщо більшість точок діаграми розсіювання знаходяться під прямою лінією, це означає, що дані, одержані за допомогою методу, є занадто високими. Середня відмінність в даному випадку буде позитивною. Нульове середнє відмінностей говорить про відсутність відхилень в об'єктивному методі. Середні відхилення для трьох об'єктивних методів показані на рис. 6. Аналіз не показує значних відхилень в, заснованому на відношенні сигнал-шум, методі, що дає очікуваний збіг кривої другого порядку (див рис. 2). І модуляційний, і %ALcons методи показують відхилення в наступних вимірюваннях, які занадто малі в порівнянні з дійсними значеннями.

Треба підкреслити, що метод, використаний для конвертації відношення сигнал-шум за Лохнер-Бургером в %ФБпред, був розроблений за допомогою даних тільки даного дослідження і не може бути узагальненим без більш докладного вивчення. Аналіз регресії був необхідний, бо не існує універсального методу, створеного попередніми дослідженнями.

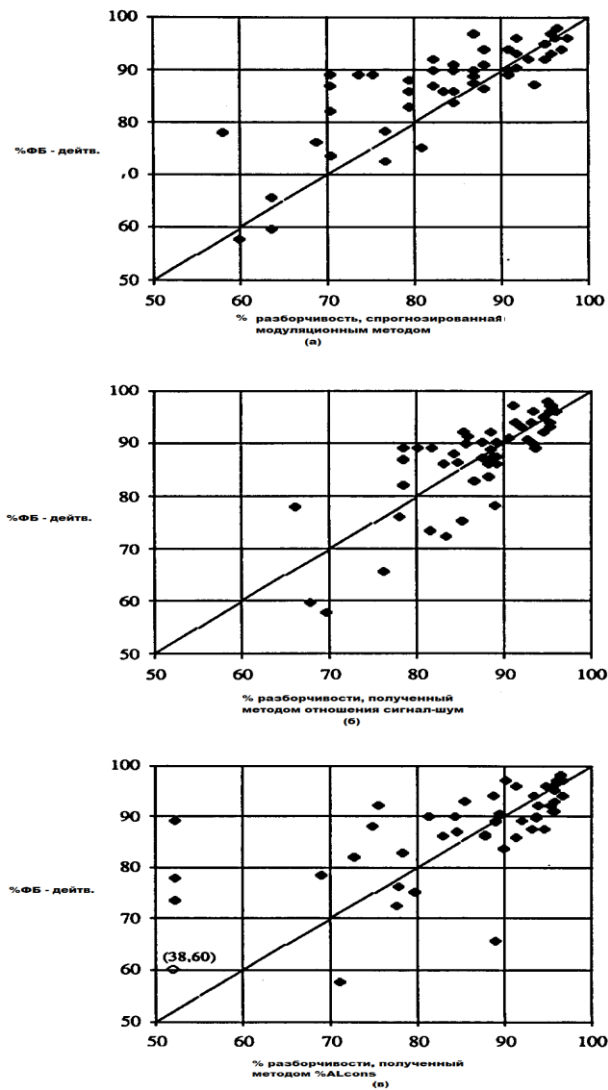


Рис. 4 Диаграммы рассеяния %ФБдейств относительно %ФБпред для трех объективных методов. (а) Модуляционный. (б) метод отношения сигнал-шум. (в) метод %ALcons

Залежність результатів в областях, найбільш цікавих для розробників

Інший шлях перевірити дані - це задатися питанням, чи є області, що представляють особливий інтерес для розробників звукових систем. У контексті створення звукових систем, їх як мінімум чотири. Яка точність трьох об'єктивних методів, якщо:

- Дійсна розбірливість - менше 85%;
- Дійсна розбірливість - більше 85%;
- Об'єктивно певна розбірливість - менше 85%;
- Об'єктивно певна розбірливість - більше 85%.

Величина 85% ФБ розбірливості була обрана, так як це загальноприйнята межа між непоганою і відмінною розбірливістю. Перше питання стосується точності об'єктивних методів у випадках, коли вже відомо, що розбірливість неприйнятна, другий - коли дійсна розбірливість достатньо хороша. Третє питання є важливим для випадків, коли об'єктивні методи показують недостатню розбірливість, і четверте - коли розбірливість, згідно з об'єктивними методами, є допустимою. Результати показані на рис. 7.

Рис.7 показує, що в кожній з чотирьох областей, що представляють особливий інтерес для розробників звукових систем, модуляційний метод і метод відношень сигнал-шум мають меншу стандартну помилку, ніж метод %ALcons. Інформація, що міститься в графіках, також показує, що об'єктивні методи найгірше працюють для області нижче порога 85%-ної розбірливості.

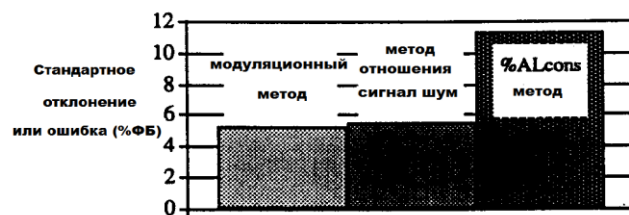


Рис. 5 Стандартное отклонение объективных методов

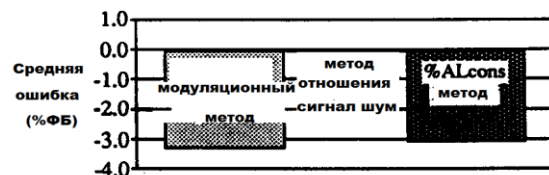


Рис. 6 Главное отличие между объективно полученными данными и действительной разборчивостью. Отрицательное среднее отличие соответствует сопоставимо малым объективно полученным значениям

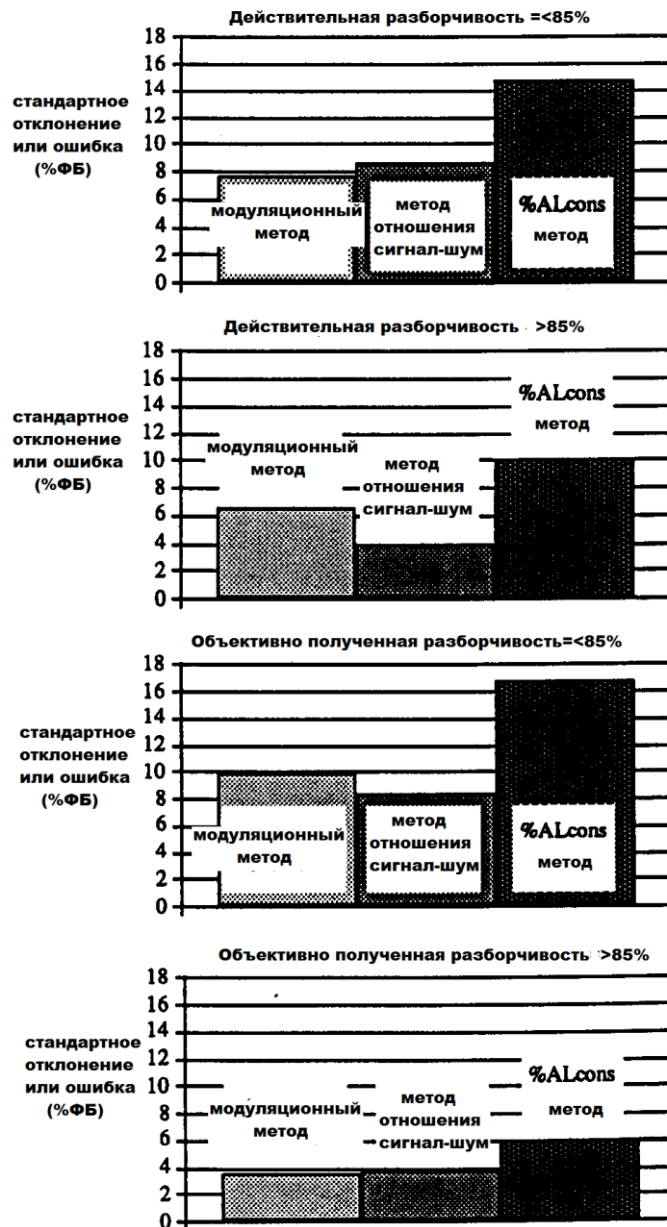


Рис. 7. Стандартные отклонения для четырех случаев, представляющих особый интерес разработчикам звуковых систем.

Илюстраційні приклади

Досліджуючи дані дійсної розбірливості (див. Додаток табл. А.1) можна виявити несподіваний результат, а саме: складний динамік сприймався як більш розбірливий, в порівнянні з «ріжковим» джерелом, що має більш вузьку спрямованість (тильна частина собору). В даному випадку модуляційний метод і метод відношення сигнал-шум пророкують однакову тенденцію серед джерел, в той час як метод %ALcons каже, що «ріжок» буде значно більш розбірливий, як показано на рис. 9.

Ці індивідуальні умови наводять на думку про необхідність включення ранніх віддзеркалень в об'єктивні методи визначення розбірливості. Більш того, даний випадок показує, що звуження спрямованості джерела призводить до появи додаткового прямого звуку, що корисно для підвищення розбірливості, проте ігнорування інших факторів, що впливають на відношення ранніх звуків до більш пізніх, може привести до серйозних помилок.

Висновки

Представлені дані підтверджують висновок про те, що індекс розбірливості, отриманий модуляційним методом, є точним показником розбірливості для «ревербераційних» приміщень. Узгодженість між дійсною і об'єктивно отриманою розбірливістю в цілому була дуже високою, і в області нижче 85% розбірливості, і в області вище. Ці результати підтверджують придатність STI, як міжнародного стандарту для об'єктивної оцінки розбірливості мови.

Метод відношення сигнал-шум Лохнера-Бургера, проявив свою точність, лише коли до отриманих даних була застосована унікальна функція для перетворення відношення сигнал шум в значення %ФБ. У будь-якому випадку, було показано, що відношення сигнал-шум містить інформацію, необхідну для оцінювання розбірливості, якщо отримані дані точні. При цьому універсального методу для конвертації відносин сигнал шум в %ФБ розбірливість запропоновано не було. Потрібні додаткові дослідження для встановлення загальної залежності між відношенням сигнал-шум Лохнера-Бургера і розбірливістю.

Було показано, що використана формула для визначення %ALcons (вираз 5), найбільш точна. З результатів можна зробити висновок, що визначення розбірливості за цією формулою може привести до серйозних помилок. На додаток, гіпотеза про те, що розбірливість може зростати монотонно, зі збільшенням спрямованості джерела, є помилковою. Важливо відзначити, що при одних і тих же умовах, модуляційний метод і метод відношення сигнал-шум

точно оцінюють розбірливість, незалежно від спрямованості джерела. Дані приводять до висновку, що спрямованість джерела важлива тільки в сенсі її впливу на відношення ранніх і пізніх звукових складових. У висновку, цей експеримент був проведений і представлений в манері, що дозволяє повторення іншими дослідниками. Дані приміщень, джерел і позицій слухача можуть бути в подальшому використані для встановлення більш точних об'єктивних методів оцінювання розбірливості.

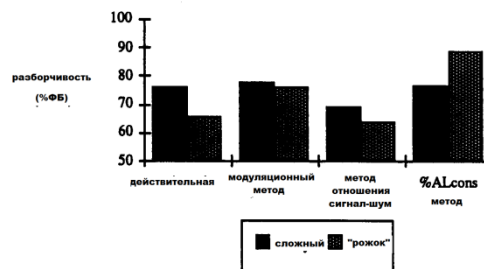


Рис. 8. Действительные значения разборчивости и полученные объективными методами в задней части собора South End. Стоит отметить, что модуляционный метод и метод отношения сигнал-шум дают результаты со схожей тенденцией изменения, в то время как метод %ALcons дает противоположные данные.

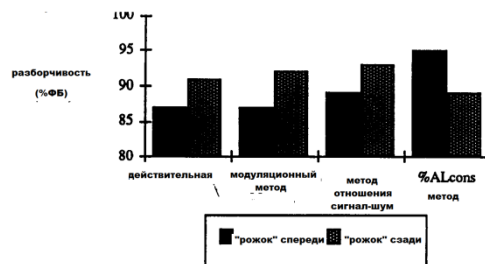


Рис. 9. Объективно полученные и действительные значения разборчивости в передней и задней частях Mechanics Hall с использованием "рожкового" источника (см. дополнение табл. А.1) Стоит отметить, что модуляционный метод и метод отношения сигнал-шум дают результаты со схожей тенденцией изменения, в то время как метод %ALcons - противоположной

Табл. А.1 Субъективные данные*

помещение	источник	позиция	значение
Berklee	Сфера	близко	96 ± 1.0
		далеко	93 ± 1.0
	Сложный	близко	96 ± 1.0
		далеко	96 ± 1.0
	"Рожок"	близко	98 ± 0.7
		далеко	96 ± 1.0
Coolidge	Сфера	близко	97 ± 0.6
		далеко	90 ± 1.6
	Сложный	близко	97 ± 0.6
		далеко	94 ± 1.1
	"Рожок"	близко	97 ± 0.6
		далеко	91 ± 1.2
Huntington	Сфера	близко	94 ± 1.3
		далеко	86 ± 1.7
	Сложный	близко	95 ± 1.0
		далеко	89 ± 1.4
	"Рожок"	близко	94 ± 1.1
		далеко	92 ± 1.1
Bridget's	Сфера	близко	92 ± 1.1
		далеко	82 ± 2.1
	Сложный	близко	92 ± 1.2
		далеко	88 ± 1.5
	"Рожок"	близко	93 ± 1.2
		далеко	86 ± 1.5
Nevins	Сфера	близко	78 ± 3.0
		далеко	89 ± 1.6
	Сложный	близко	87 ± 3.0
		далеко	89 ± 1.1
	"Рожок"	близко	89 ± 1.6
		далеко	90 ± 1.6
Jordan	Сложный	близко	89 ± 1.4
		далеко	78 ± 2.4
	"Рожок"	близко	90 ± 1.9
		далеко	87 ± 1.6
Mechanic's	Сложный	близко	86 ± 2.0
		далеко	83 ± 2.0
	"Рожок"	близко	87 ± 2.0
		далеко	91 ± 1.6
Cathedral	Сложный	близко	90 ± 1.8
		далеко	76 ± 3.7
	"Рожок"	близко	91 ± 1.7
		далеко	66 ± 2.6
Cyclorama	Сложный	близко	86 ± 2.6
		далеко	73 ± 2.3
	"Рожок"	близко	87 ± 2.3
		далеко	72 ± 2.9
MIT Track	Сложный	близко	75 ± 2.5
		далеко	60 ± 3.2
	"Рожок"	близко	84 ± 2.9
		далеко	58 ± 3.4

* Помещение - аудитории перечисленные в табл. 2; Источник - динамики из табл 1; Позиция - позиция слушателя, определенная в 1.4; Значение - среднее по всем субъектам, с точностью сопоставимой с реальной разборчивостью, рассчитанной с использованием теста Student-ca.

Табл. А.2. Индексы рече-передачи, отношения сигнал-шум и значения %ALcons для всех условий*

помещение	Источник	Позиция	STI	С/Ш отношение	%ALcons pred
Berkeley	Sphere	Near	0.65	6.6	8
		Far	0.71	8.9	8
	Array	Near	0.72	8.8	1
		Far	0.72	8.9	5
	Horn	Near	0.73	8.6	1
Coolidge	Sphere	Far	0.78	11.2	2
		Near	0.60	5.0	5
	Array	Far	0.56	3.2	9
		Near	0.71	9.2	1
	Horn	Far	0.64	6.4	3
Huntington	Sphere	Near	0.71	8.7	0
		Far	0.61	4.7	1
	Array	Near	0.61	5.2	6
		Far	0.57	3.5	10
	Horn	Near	0.70	7.8	1
Bridget's	Sphere	Far	0.64	6.9	4
		Near	0.74	9.0	0
	Array	Far	0.67	7.8	1
		Near	0.56	2.2	17
	Horn	Far	0.48	-0.3	17
Nevins	Sphere	Near	0.70	3.6	3
		Far	0.54	1.7	17
	Array	Near	0.65	5.6	1
		Far	0.54	1.3	7
	Horn	Near	0.41	-3.8	32
Jordan	Sphere	Far	0.48	-0.3	32
		Near	0.48	-0.3	11
	Array	Far	0.51	0.8	32
		Near	0.50	0.2	4
	Horn	Far	0.60	4.0	12
Mechanic's	Sphere	Near	0.60	3.6	7
		Far	0.52	3.8	20
	Array	Near	0.64	6.6	2
		Far	0.56	3.1	9
	Horn	Near	0.58	3.9	5
Cathedral	Sphere	Far	0.54	2.7	16
		Near	0.60	4.0	2
	Array	Far	0.65	6.1	6
		Near	0.58	2.3	3
	Horn	Far	0.47	-0.5	17
Cyclorama	Sphere	Near	0.58	2.4	1
		Far	0.44	-1.1	6
	Array	Near	0.61	1.9	8
		Far	0.48	0.7	32
	Horn	Near	0.68	3.7	3
MIT Track	Sphere	Far	0.52	1.4	14
		Near	0.55	2.1	15
	Array	Far	0.44	-3.4	41
		Near	0.58	3.5	5
	Horn	Far	0.42	-2.9	19

*STI - индекс рече-передачи, полученный из модуляционной передаточной функции, как описано в 5.1; С/Ш отношение - измеренное отношение сигнал-шум Лохнера-Бургера, как описано в 5.2; %ALcons pred - процент артикуляционно-потерянных согласных как описано в 5.3

Лекція 1.9. Програмне забезпечення та апаратурне обладнання

1.9.1. Коригування частотної характеристики вимірювального тракту

В даному підрозділі представлено технологію коригування частотної характеристики вимірювального тракту апаратно-програмного комплексу «Штучна голова», призначеного для акустичної експертизи приміщень шляхом об'єктивного оцінювання розбірливості мови. Показано, що таке коригування може бути виконано шляхом контрольованого ділення частотної характеристики системи «гучномовець-приміщення-мікрофон» на попередньо отриману оцінку амплітудної частотної характеристики підсистеми «гучномовець-мікрофон». Виконано аналіз характеру та ступеня впливу такого коригування на точність оцінювання імпульсної характеристики приміщення. Показано необхідність застосування методу регуляризації для досягнення задовільної точності обчислень, вказано можливий варіант регуляризуючого множника та значення параметра регуляризації.

Вступ

Штучну голову (ШГ), що містить модель зовнішнього вуха, використовують в акустиці досить давно при розв'язанні таких задач як запис звуку в концертній залі, запис відгуку приміщення на звуковий стимул для подальшого обчислення імпульсної характеристики (ІХ) приміщення, контроль якості навушників, контроль якості та розбірливості мови в лініях зв'язку [1-3].

Прикладом ефективного використання штучної голови є база даних, що містить записи бінауральних імпульсних характеристик кількох приміщень різного призначення [4], [5], [7]. Науково-практична цінність таких записів полягає в можливості моделювання звуку в різних точках приміщення (ефект «акустичне дзеркало» [8]), оцінювання якості та розбірливості мовних презентацій, якості музичних презентацій, дослідження алгоритмів дереверберації [9], [10], [11], [12], [13], [14].

На жаль, згадана база бінауральних ІХ [4], [5] є досить обмеженою. Наприклад, вона містить записи бінауральних ІХ лише однієї аудиторії, тоді як архітекторам та звукоінженерам корисно було б мати бінауральні ІХ для низки лекційних та офісних приміщень різних розмірів та із різними архітектурними особливостями. Для усунення вказаного недоліку потрібно мати апаратно-програмний комплекс «Штучна голова», із використанням якого можна отримати низку акустичних портретів у вигляді оцінок бінауральних ІХ для приміщень різного призначення. Далі обговорюються питання створення апаратно-програмного комплексу саме такого типу.

Постановка задачі

Спрощену структурну схему процедури оцінювання ІХ приміщення за допомогою ШГ представлено на рис. 1.



Рис. 1. Структурна схема процедури оцінювання ІХ приміщення

Тестовий звуковий сигнал $x(t)$ випромінюють за допомогою гучномовця, розміщеного в точці приміщення, де зазвичай знаходиться диктор. Відгук приміщення $y(t)$ на стимул $x(t)$ сприймають парою мікрофонів, притиснутих до вушних раковин штучної голови, розміщеної в одній із точок приміщення, де зазвичай знаходяться слухачі.

Оскільки амплітудно-частотні характеристики гучномовця та мікрофона не є ідеально рівномірними в смузі частот тестового сигналу, замість ІХ приміщення $h_r(t)$ насправді буде оцінено згортку

$$h_{\Sigma}(t) = h_l(t) \otimes h_r(t) \otimes h_m(t) = h_r(t) \otimes h_{lm}(t), \quad (1)$$

де \otimes - символ згортки; $h_l(t)$ - ІХ гучномовця; $h_m(t)$ - ІХ одного із мікрофонів; $h_{lm}(t) = h_l(t) \otimes h_m(t)$ - ІХ системи «гучномовець-мікрофон» (ГМ). $h_\Sigma(t)$

Спосіб обчислення ІХ приміщення $h_r(t)$ є досить очевидним. Дійсно, оскільки в частотній області співвідношенню (1) відповідає добуток відповідних частотних характеристик

$$H_\Sigma(f) = H_l(f) \cdot H_r(f) \cdot H_m(f) = H_r(f) \cdot H_{lm}(f), \quad (2)$$

де $H(f) = \mathbb{F}\{h(t)\}$, \mathbb{F} - символ прямого перетворення Фур'є, із (2) отримуємо

$$h_r(t) = \mathbb{F}^{-1}\{H_r(f)\} = \mathbb{F}^{-1}\left\{\frac{|H_\Sigma(f)|}{|H_{lm}(f)|} \cdot \exp[\theta_\Sigma(f) - \theta_{lm}(f)]\right\}, \quad (3)$$

де $\theta_\Sigma(f)$ та $\theta_{lm}(f)$ - фазочастотні характеристики (ФЧХ) функцій $H_\Sigma(f)$ та $H_{lm}(f)$, відповідно.

Проблемним місцем для реалізації обчислень за виразом (3) є операція ділення, оскільки модуль $|H_{lm}(f)|$ (амплітудно частотна характеристика (АЧХ) підсистеми ГМ) функції $H_{lm}(f)$ може містити малі числові значення, що призведе до переповнення розрядної сітки обчислювальної системи та аварійної зупинки комп'ютера. Разом із тим, зрозуміло, що якщо забезпечити належний контроль за властивостями функції $|H_{lm}(f)|$, рівняння (3) може використовуватися для обчислень.

Що стосується ФЧХ $\theta_\Sigma(f)$ та $\theta_{lm}(f)$, тут є два варіанти: 1) враховувати інформацію про обидві ФЧХ; 2) враховувати інформацію лише про $\theta_\Sigma(f)$, припускаючи, що $\theta_{lm}(f) = 0$.

Таким чином, метою даної роботи є пошук методів обчислення АЧХ підсистеми «гучномовець-мікрофон», достатньо точних та відносно простих з точки зору їх технічної реалізації.

Організація досліджень

Доцільно скористатися інформацією, викладеною в [4] щодо властивостей апаратно-програмного комплексу, що був використаний при оцінюванні бінауральних ІХ різних приміщень.

Загалом, вимірювання проводилися за допомогою професійної аудіоапаратури, щоб отримати високоякісні та низькошумні результати. В якості джерела звуку було використано акустичний двоканальний активний студійний монітор Genelec 8130, що має цифровий інтерфейс AES/EBU. Для запису звуку були використані два всенаправлених конденсаторних вимірювальних мікрофони Beyerdynamic MM1 та мікрофонний підсилювач RME Octamic II у поєднанні з аудіоінтерфейсом RME Multiface II. Всі вимірювання проводилися з частотою дискретизації 48 кГц і 24-бітовою точністю. Для бінауральних записів використано штучну голову HMS2 фірми «HEAD acoustics». Два мікрофони розташовувалися впритул до вушної раковини на відстані 1 см від вушного каналу.

Базовим елементом при побудові тестового сигналу слугував mls-сигнал із ступенем 16 (тобто сигнал містив 2^{16} вибірок), що відповідає довжині сигналу 1,36 с при частоті дискретизації 48 кГц. Тестовий сигнал створювався шляхом повторювання цього базового елемента 17 разів. Імпульсну характеристику приміщення обчислювали, усереднюючи 16 останніх сплесків оцінки взаємнокореляційної функції (ВКФ) між сигналом з виходу мікрофона та тестовим сигналом, що дозволило підвищити відношення сигнал-шум на 12 дБ.

Автори даної статті планували використовувати такі ж тестові сигнали, як і в [4], з тією лише відмінністю, що частота дискретизації становила 44,1 кГц. Якщо вважати ланцюжок «гучномовець-приміщення-мікрофон» лінійною системою із ІХ $h_{\Sigma}(t)$, тоді, як відомо [5], така ІХ може бути обчисленою за виразом

$$K_{xy}(\tau) = \langle x(t)y(t+\tau) \rangle = \int_{-\infty}^{\infty} h_{\Sigma}(v) \langle x(t)x(t+\tau-v) \rangle dv = \int_{-\infty}^{\infty} h_{\Sigma}(v) K_x(\tau-v) dv \quad (4)$$

де $\langle \cdot \rangle$ - символ математичного очікування; $K_x(\tau)$ - автокореляційна функція тестового сигналу $x(t)$; $K_{xy}(\tau)$ - ВКФ тестового сигналу $x(t)$ та сигналу $y(t)$ на виході мікрофона. Дійсно, за умови

$$K_x(\tau) = D_x \cdot \delta(\tau),$$

де $\delta(\tau)$ - дельта-функція Дірака, D_x - дисперсія тестового сигналу $x(t)$, із (2) отримуємо

$$K_{xy}(\tau) = D_x \cdot h_\Sigma(\tau). \quad (5)$$

Фур'є-перетворення від (5) приводить до виразу

$$S_{xy}(f) = H_\Sigma(f) \frac{N_0}{2}, \quad (6)$$

де $S_{xy}(f)$ - взаємний спектр сигналів $x(t)$ та $y(t)$; N_0 - спектральна щільність білого шуму $x(t)$.

Якщо вимірювання виконувати в заглушеній кімнаті, на малій відстані між мікрофоном та гучномовцем, тоді, згідно (5) та (6), можна отримати інформацію про ІХ та частотну характеристику підсистеми ГМ:

$$K_{xya}(\tau) = D_x \cdot h_{lm}(\tau), \quad S_{xya}(f) = H_{lm}(f) \frac{N_0}{2}, \quad (7)$$

оскільки форма вихідного сигналу $y(t)$ буде визначатися властивостями підсистеми ГМ.

Оскільки при практичних вимірюваннях запис та обробка сигналів виконуються у цифровій формі, при використанні виразів (7) достатньо забезпечити, щоб спектр потужності тестового сигналу $x(t)$ був рівномірним у смузі частот $F_s/2$, де F_s - частота дискретизації. Для обраного тестового mls-сигналу можна вважати, що ця вимога виконується із достатньою для практичної діяльності точністю, оскільки ширина центрального сплеску є мінімально можливою та становить $1/F_s$, а рівень бокових пелюстків є близьким до мінус 47 дБ (рис. 2).

Приклади форми сигналів $x(t)$ та $y(t)$ на вході та виході, відповідно, підсистеми ГМ наведено на рис. 3, а відповідних спектрів потужності – на рис. 4. Як бачимо із рис. 3, вибірки сигналу $y(t)$, на відміну від вибірок сигналу $x(t)$, мають різну висоту, що свідчить про спотворення вхідного сигналу $x(t)$ підсистемою ГМ. Наведені на рис. 4 оцінки спектрів потужності більш інформативні, оскільки дозволяють визначити, на яких частотах відбувається найбільш сильне спотворення вхідного сигналу. Дійсно, якщо нерівномірність спектру вхідного сигналу (рис. 4а) не перевищує $\pm 0,5$ дБ, то нерівномірність спектру вихідного сигналу в смузі частот від 300 Гц до 10 кГц становить ± 3 дБ, є близькою до ± 6 дБ в смузі частот від 10 кГц до 16 кГц, а в смузі від 50 Гц до 300 Гц спостерігається провал глибиною мінус 15 дБ.

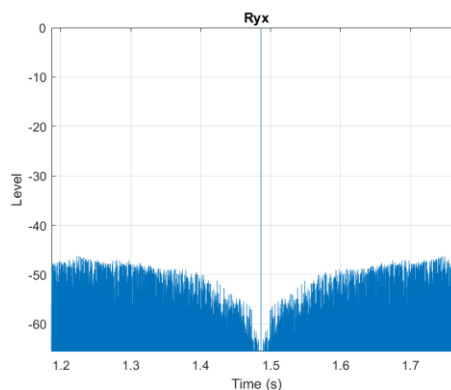


Рис.2. Автокореляційна функція mls-сигналу

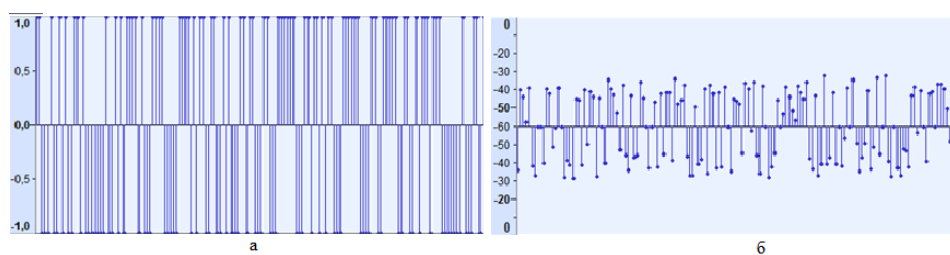


Рис. 3. Сигнали $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ

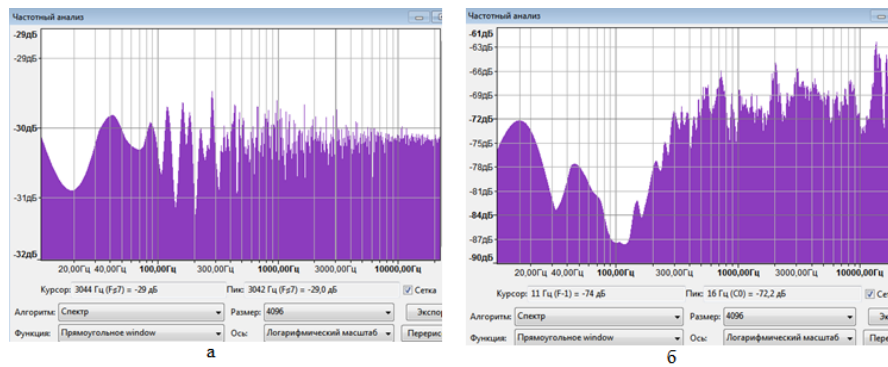


Рис. 4. Спектри потужності сигналів $x(t)$ (а) та $y(t)$ (б) на вході та виході підсистеми ГМ

Згідно із (7), форму АЧХ підсистеми можна також визначити шляхом взаємно-кореляційного аналізу сигналів $x(t)$ та $y(t)$. На рис. 5 наведено форму оцінки ВКФ

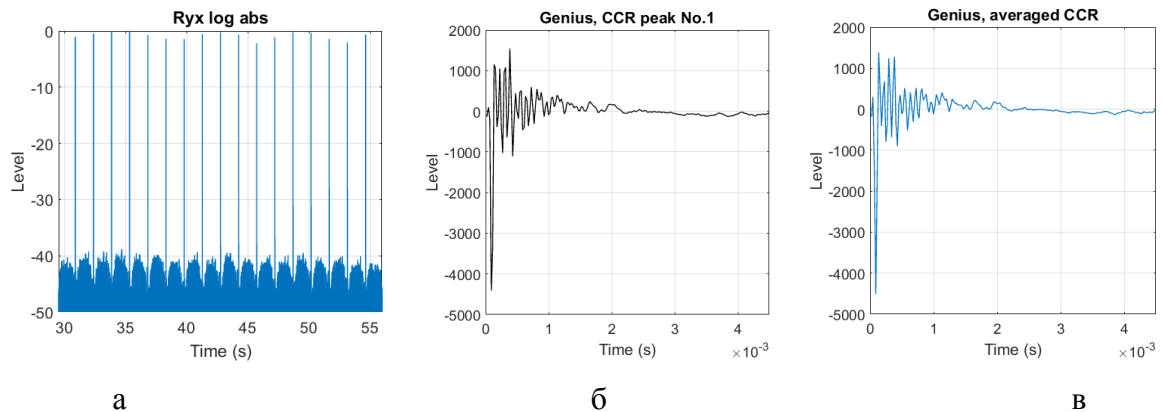


Рис. 5. ВКФ сигналів $x(t)$ та $y(t)$ (а), форма першого (б) та усередненого (в) сплесків ВКФ

Модуль Фур'є-перетворення усередненого сплеску ВКФ, згідно (7), є оцінкою АЧХ підсистеми ГМ.

Результати досліджень

На рис. показано оцінки АЧХ $|H_{lm}(f)|$ підсистеми ГМ для 3-х видів гучномовців:

- 1) малогабаритний MP3-плеєр Optima МК-4 (вихідна потужність 5 Вт, частотний діапазон 150-18000 Гц, відношення сигнал-шум 86 дБ);
- 2) активна колонка Genius SP-HF 2.0 500 (вихідна потужність 14 Вт, частотний діапазон 65-20000 Гц, відношення сигнал-шум 73 дБ);

3) активні акустичні монітори Yamaha MSP5 (вихідна потужність 67 Вт, частотний діапазон 50-40000 Гц, відношення сигнал-шум 102 дБ).

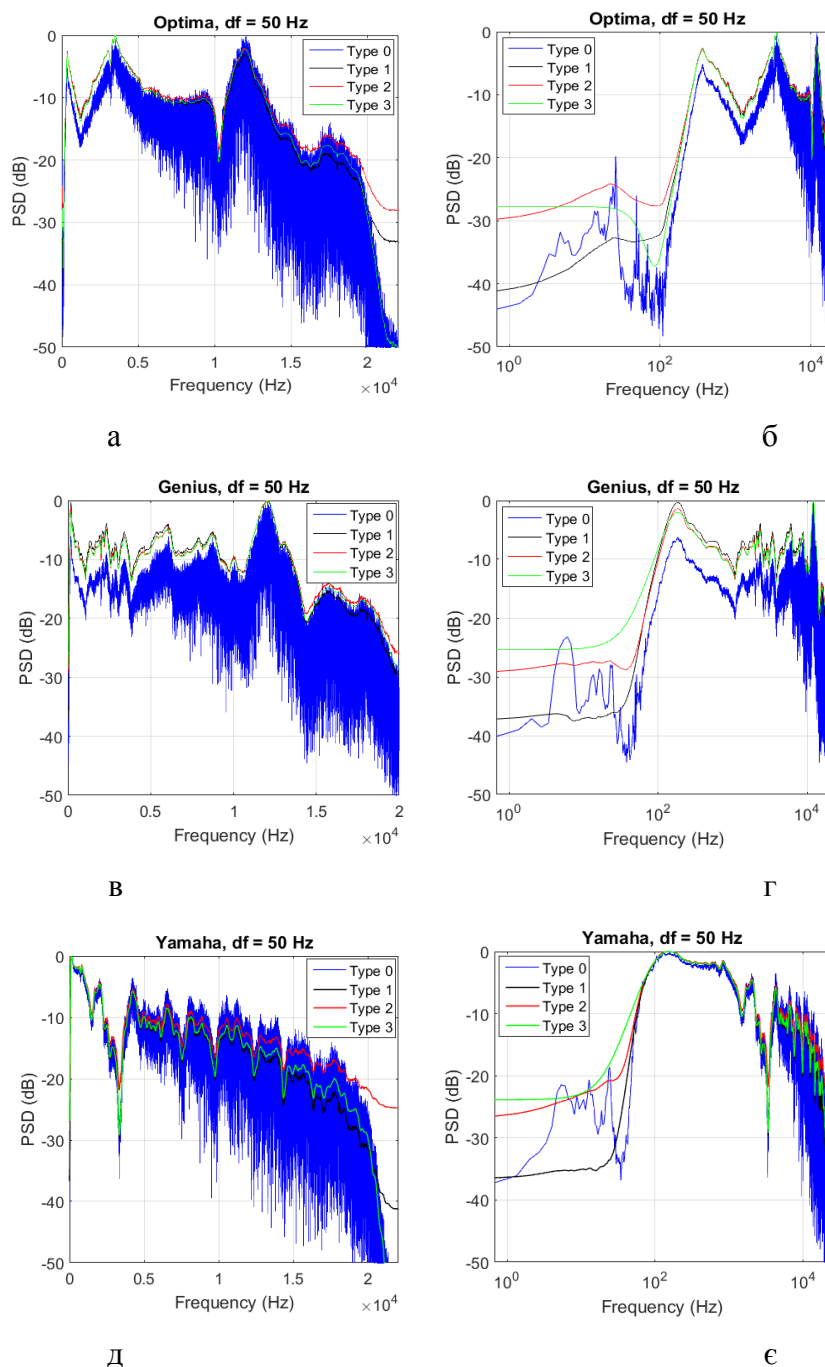


Рис. 6. Оцінки спектру сигналу гучномовців (заглушене приміщення)

На кожному із рисунків наведено графіки оцінок АЧХ $|H_{lm}(f)|$ чотирьох типів:

1) Type 0 - оцінка спектра потужності Велча як усереднення спектрів потужності 17 сегментів сигналу $y(t)$, кожен сегмент має довжину 65536 вибірок

та не піддається ваговому зважуванню:

$[Py, f_{yy}] = \text{pwelch}(y, \text{rectwin}(65536), 0, 65536, F_s);$

2) Type 1 - згладжений за частотою, трикутним вікном шириною 50 Гц, модуль дискретного перетворення Фур'є (ДПФ) від результату усереднення 17 піків ВКФ;

3) Type 2 - згладжений за частотою, трикутним вікном шириною 50 Гц, результат усереднення (17 разів) модулів ДПФ від 17 сплесків ВКФ;

4) Type 3 - оцінка спектра потужності Велча із використанням вагового вікна Ханна, що забезпечує роздільну здатність 50 Гц: $[P_{yy}, f_{yy}] = \text{pwelch}(y, \text{hann}(\text{delw}), \text{round}(\text{delw}/2), 65536, F_s), \text{delw} = (1, 5F_s)/df, F_s = 44100 \text{ Гц}, df = 50 \text{ Гц}.$

Процедуру згладжування трикутним вікном, використану в оцінках 1-го та 2-го типів, було реалізовано шляхом цифрової фільтрації незгладженої оцінки спектру, для чого було застосовано рекурсивний фільтр першого порядку. Це двоетапна процедура, на першому етапі якої виконується згладжування прямокутним вікном зліва направо:

$$y_n = y_{n-1} + \frac{1}{N}(x_n - x_{n-N}), \quad n = N+1, \dots, N_{fft}/2+1,$$

а на другому етапі виконується згладжування прямокутним вікном справа наліво:

$$y_n = y_{n+1} + \frac{1}{N}(x_n - x_{n+N}), \quad n = N_{fft}/2 - N + 1, \dots, 1,$$

де $N = \frac{N_{fft}}{F_s} \cdot \frac{df}{2}, N_{fft} = 65536, F_s = 44100 \text{ Гц}, df = 50 \text{ Гц}.$

Як бачимо, значення параметру N обчислювалося, виходячи із ширини прямокутного вікна 25 Гц. Це зумовлене тим, що подвійне згладжування прямокутним вікном еквівалентне одиночному згладжуванню трикутним вікном подвоєної ширини. Необхідність в застосуванні двох напрямків фільтрації зумовлена тим, що при цьому не треба дописувати якісь вибірки поза краями первинної оцінки спектру, що підлягає згладжуванню.

Порівнюючи між собою наведені на рис. 6 оцінки $|H_{lm}(f)|$, доходимо наступних висновків.

По-перше, виходячи із ступеня фактичної нерівномірності та фактичному частотному діапазону оцінки АЧХ $|H_{lm}(f)|$, слід віддати перевагу активній колонці Genius. Як це не дивно, але акустичний монітор Yamaha виявився неконкурентним, що можна пояснити його фізичним старінням.

По-друге, як видно із наведених графіків, незгладжену оцінку $|H_{lm}(f)|$ нульового типу недоцільно використовувати для подальшої корекції, з огляду на неприйнятно велику дисперсію в області середніх та, що особливо, високих частот.

По-третє, серед згладжених оцінок найбільш привабливою виглядає оцінка третього типу як така, що найкращим чином забезпечує зниження дисперсії оцінки без надмірного підвищення її зміщення. Зауважимо, що певна доля суб'єктивності при виборі типу оцінки є неминучою, оскільки істинна АЧХ підсистеми ГМ невідома.

Результати обчислень ІХ підсистеми ГМ згідно рівняння (3) із використанням згладженої оцінки $|H_{lm}(f)|$ для гучномовця Genius показані на рис. 7 для різних значень часового масштабу.

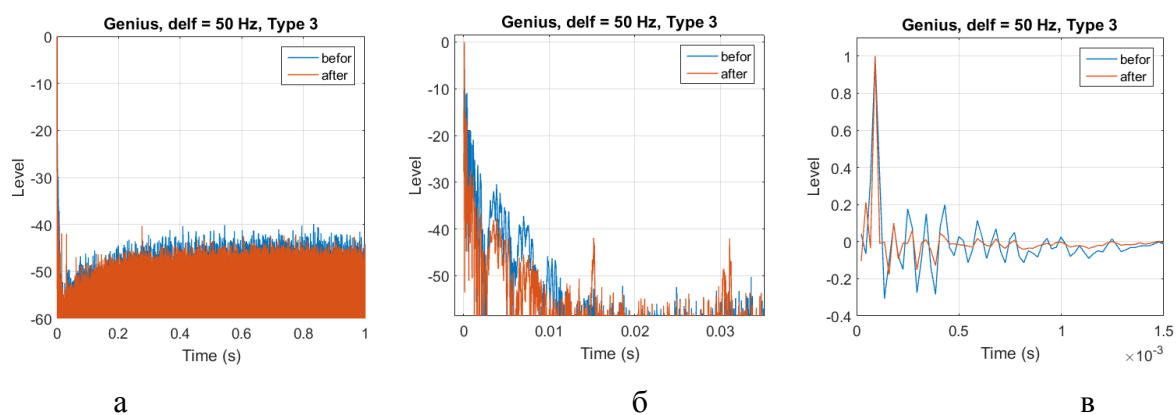


Рис. 7. Скориговані оцінки ІХ для гучномовця Genius

Результат коригування, як бачимо, має як позитивні, так і негативні якості. Негативною є поява окремих дискретних сплесків достатньо високого рівня -40

дБ (рис. 7а). Позитивні: 1) рівень бокових пелюстків знизився приблизно на 3-4 дБ (рис. 5а); 2) головний пік ІХ став вужчим в 2 рази, а рівень найближчих до нього сплесків став нижчим майже на 6 дБ (рис. 5б,в).

З огляду на гостроту зазначених вище сплесків доречно припустити, що вони спричинені збільшенням дисперсії оцінки АЧХ, отриманої із використанням виразу (6), на високих частотах, через ділення на малі значення згладженої оцінки $|H_{lm}(f)|$. Очевидним та простим способом пригнічення таких сплесків є множення відкоригованої оцінки АЧХ на вагове вікно. В часовій області це відповідає процедурі згладжування оцінки ІХ шляхом ковзного усереднення.

В рамках даної роботи розглянуто два види такого вагового вікна. В першому випадку це вікно описується виразом:

$$w_1(r) = \begin{cases} 1, & r = 1, \dots, \frac{N_{fft}}{2} - k, \quad r = \frac{N_{fft}}{2} + 2 + k, \dots, N_{fft}, \\ 0,5 \left[1 - \cos \frac{\pi}{k} \left(r - \frac{N_{fft}}{2} - 1 \right) \right], & r = \frac{N_{fft}}{2} + 1 - k, \dots, \frac{N_{fft}}{2} + 1 + k, \end{cases}$$

де k визначає інтервал, що відповідає половині періоду косинусоїди, що описує кінцеву частину вікна, яка відповідає за пригнічення високих частот:

$$k = (F_s/2 - f_1) / (F_s/N_{fft}).$$

Друге вагове вікно – це вікно Ханна

$$w_2(r) = 0,5 \left[1 - \cos \left(\pi (r - N_{fft}/2 - 1) / \Delta N \right) \right], \quad r = 1, \dots, N_{fft},$$

де ΔN - половина ширини вікна Ханна у вибірках.

Графіки вагових вікон для $f_1 = 19$ кГц та $f_2 = 18$ кГц ($f_2 = \Delta N \cdot F_s / N_{fft}$) наведено на рис. 8. Неважко помітити, що в окремому випадку $f_1 = 0$, $f_2 = F_s/2$ ці вікна співпадають: $w_1(r) = w_2(r)$.

Результати застосування вікон $w_1(r)$ та $w_2(r)$ наведено на рис. 9 та 10.

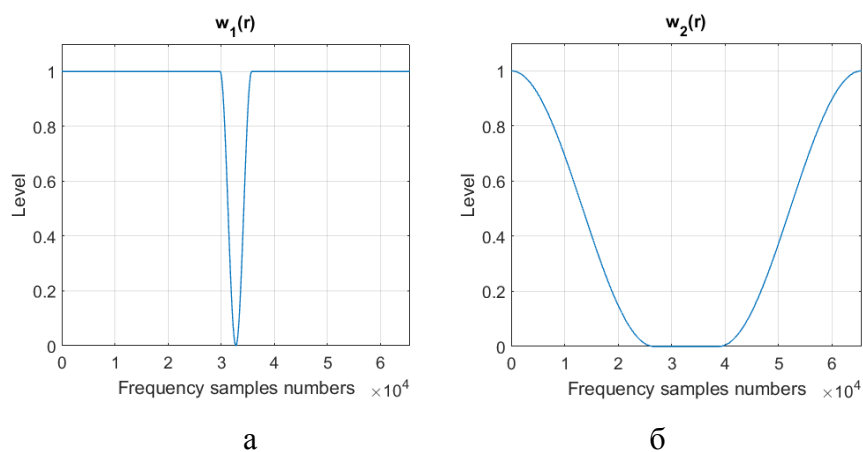


Рис. 8. Вагове вікно $w_1(r)$ для $f_1=19$ кГц (а) та вікно $w_2(r)$ для $f_2=18$ кГц (б)

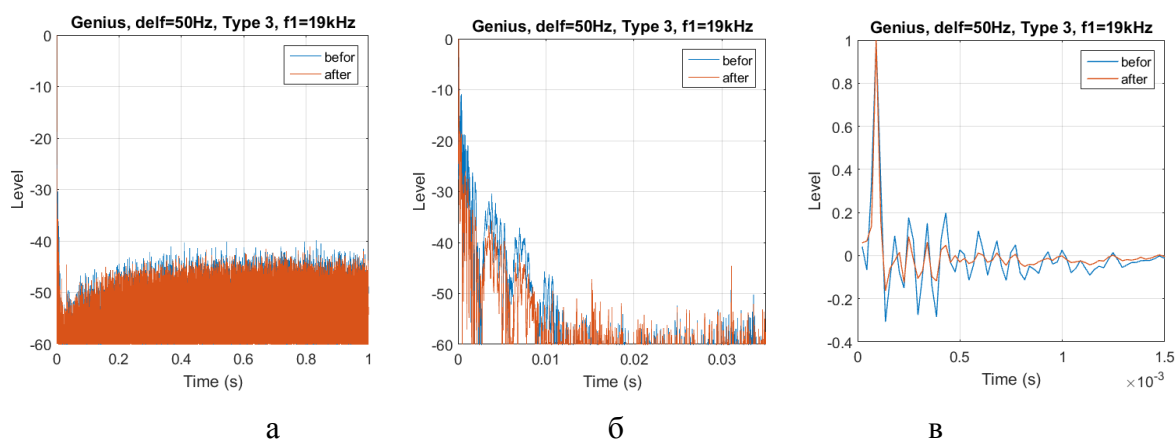


Рис. 9. Скориговані ІХ для гучномовця Genius, із вікном $w_1(r)$ для $f_1=19$ кГц

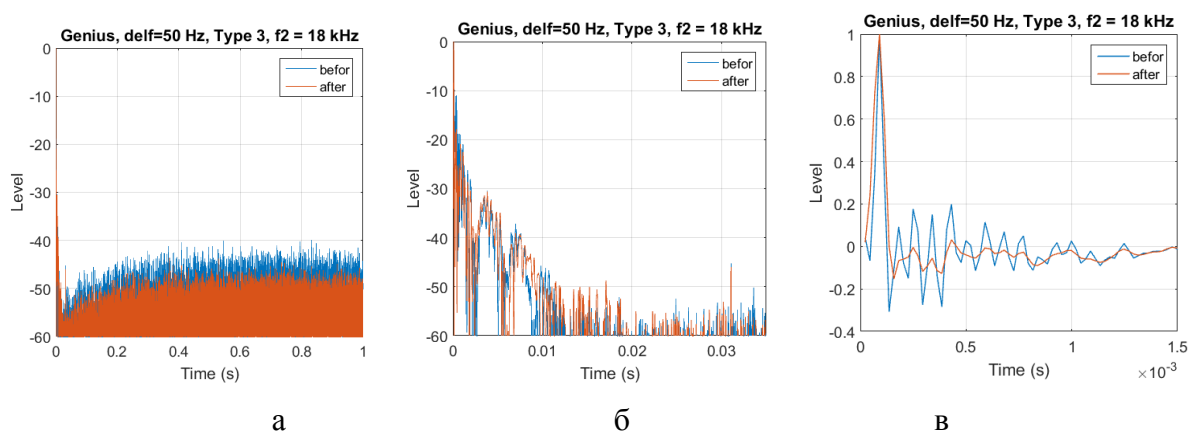


Рис. 10. Скориговані ІХ для гучномовця Genius, із вікном $w_2(r)$ для $f_2=18$ кГц

Порівнюючи рис. 9 та 10, неважко зробити висновок, що вікно $w_2(r)$ із параметром $f_2=18$ кГц є більш перспективним, оскільки дозволяє зменшити рівень бокових пелюстівок оцінки ІХ підсистеми ГМ на 4 дБ, тоді як вікно $w_1(r)$ із

параметром $f_1=19$ кГц дозволяє зменшити рівень бокових пелюстків лише на 1 дБ (рис. 9а та 10а). Правда, в інтервалі 0-10 мс вікно $w_1(r)$ є кращим, оскільки забезпечує рівень бокових пелюстків на 6 дБ нижчий, ніж вікно $w_2(r)$. Проте ця перевага не є принциповою, оскільки у випадку вікна $w_2(r)$ рівень цих бокових пелюстків не перевищує -30 дБ, а ширина головного сплеску ІХ на рівні -30 дБ не перевищує 2 мс, що забезпечує достатню для інженерних застосувань точність оцінювання ІХ приміщення.

Зазначимо, що використання вагового вікна $w_2(r)$ можна розглядати як регуляризацію, із регуляризуючим множником $w_2(r)$ та параметром регуляризації f_2 , рішення нестійкої оберненої задачі [15].

Оскільки в даній роботі розглянуто варіант коригування частотної характеристики тракту ШГ, де враховується лише нерівномірність АЧХ підсистеми ГМ, в подальшому слід розглянути доцільність додаткового врахування оцінки фазової характеристики підсистеми ГМ.

Висновки

Розроблено технологію коригування частотної характеристики апаратно-програмного комплексу «Штучна голова». Показано, що таке коригування може бути виконано на етапі обробки вихідних сигналів мікрофонів шляхом контрольованого ділення оцінки частотної характеристики системи «гучномовець-приміщення-мікрофон» на попередньо отриману оцінку АЧХ підсистеми «гучномовець-мікрофон». Виконано аналіз характеру та ступеня впливу такого коригування на точність оцінювання імпульсної характеристики приміщення та показано, що задовільну для практичного застосування точність оцінювання можна забезпечити за умови регуляризації рішення. Запропоновано рекомендації з вибору регуляризуючого множника та параметра регуляризації.

Література до підрозділу 1.7.2.1

1. J. Blauert Ed., The technology of binaural listening. Springer, Berlin–Heidelberg–New York, 2013.
2. HMS II.3 Artificial Head Measurement System with Ear Simulator and Artificial Mouth, Data Sheet, HEAD Acoustics GmbH, 2015.
3. P. Søndergaard and P. Majdak. The auditory modeling toolbox. In J. Blauert, editor, The technology of binaural listening, chapter 2. Springer, Berlin-Heidelberg-New York NY, 2013.
4. M. Jeub, M. Schafer, P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," Proc. of 2009 16th International Conference on Digital Signal Processing, 5-7 July 2009, Santorini-Hellas, Greece. DOI: 10.1109/ICDSP.2009.5201259.
5. Aachen Impulse Response Database. Available on-line: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>
6. M. Jeub, M. Schäfer, H. Krüger, C. M. Nelke, C. Beaugeant, and P. Vary, "Do We Need Dereverberation for Hand-Held Telephony?," Proc. of 20th International Congress on Acoustics (ICA), pp. 1-7, August 2010, Sydney, Australia.
7. G. Soulodre, N. Popplewell, and J. Bradley, "Combined effects of early reflections and background noise on speech intelligibility," Journal of Sound and Vibration, vol. 135, No.1, pp. 123-133, 1989.
8. Sound Forge Audio Studio. Available on-line: <https://www.magix.com/ru/muzyka/sound-forge/sound-forge-audio-studio/>
9. E. Habets, N. Gaubitch, and P. Naylor, "Temporal selective dereverberation of noisy speech using one microphone," Proc. of 2008 IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 4577-4580, March-April 2008.
10. J. Bradley, H. Sato, and M. Picard, "On the importance of early reflections for speech in rooms," J. Acoust. Soc. Am., vol. 113, no. 6, pp. 3233-3244, June 2003. DOI: 10.1121/1.1570439.
11. I. Arweiler, J. Buchholz, and T. Dau, "Speech intelligibility enhancement by early reflections," Proc. of 2nd Int. Symposium on Auditory and Audiological Research (ISAAR 2009), Elsinore, Denmark, August 2009. ISBN: 87-990013-2-2

12.H. Sato and J. Bradley, "Evaluation of acoustical conditions for speech communication in working elementary school classrooms," J. Acoust. Soc. Am. 106 (4), Pt. 1, pp. 2064–2077, 2004. DOI: 10.1121/1.2839283.

13.A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, D. Motorniuk, and A. Khrapachevskyi, "Objective and Subjective Assessment of the Quality and Intelligibility of Noised Speech," Proc. of 2018 IEEE International Scientific-Practical Conference "Problems of Infocommunications. Science and Technology" (PIC S&T`2018), pp. 71-74, October 9-12, 2018, Kharkiv, Ukraine.

14.A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, and D. Motorniuk, "Automated Subjective Assessment of Speech Intelligibility under Diotic and Dichotic Listening," in "Data-Centric Business and Applications," Springer, 2020 (in publishing).

15.Тихонов А. Н. О некорректных задачах линейной алгебры и устойчивом методе их решения. ДАН СССР, 1965, т. 163, № 3, с. 591—594.

Розділ 2. Методи та системи корекції акустики приміщень

Існує щонайменше 2 основних підходи до корекції акустики приміщень: пасивний та активний. В першому випадку відношеннями енергій ранніх та пізніх відбиттів керують шляхом певного вибору орієнтації та покриття відбиваючих поверхонь. В другому випадку використовують електроакустичні апаратно-програмні системи.

Лекція 2.1. Методи пасивної корекції акустики приміщень

Методи пасивної корекції акустики навчальних приміщень досить добре відображені в низці відповідних рекомендацій та стандартів [1–8].

Рекомендоване розташування поглинаючих поверхонь, рекомендоване в [1, 2], показано на рис. 1.

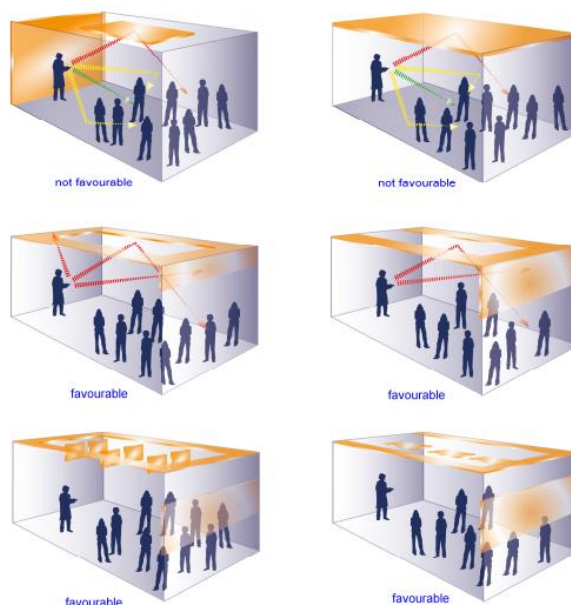


Figure 1: favourable/unfavourable distribution of sound absorption areas in a room

Цікавою особливістю розташування поглинальних поверхонь є те, що вони сконцентровані біля кутів та мають нейтралізувати підвищення тиску в зазначених місцях. Теоретично це явище обґрунтовано в роботі Ватерхауза [9], де показано, що біля відбиваючої поверхні тиск зростає на 3 дБ, біля стику 2-х стін – на 6 дБ, в кутках кімнати – на 9 дБ (Fig. 6).

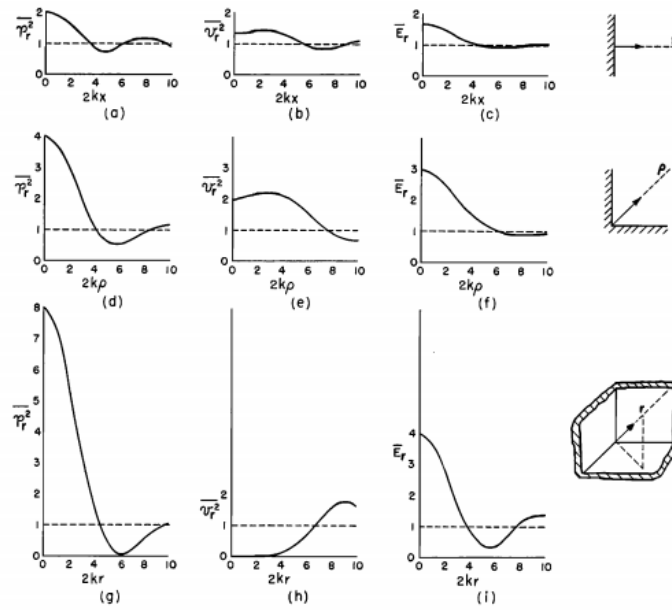


FIG. 6. Interference patterns in reverberant sound fields at reflecting walls, edges, and corners. See Table II for the expressions plotted here.

Рекомендації до час реверберації викладено на графіках Figures 2, 3 [1].

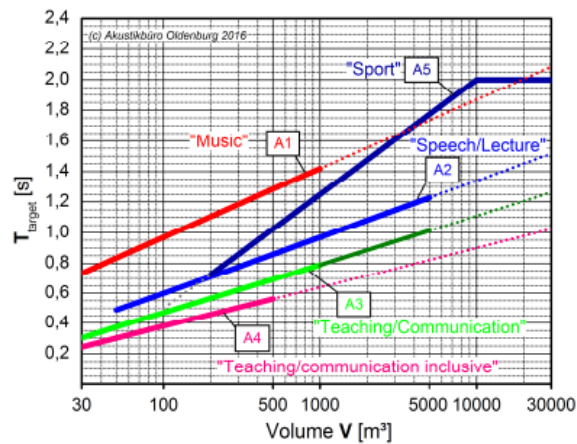


Figure 2: Values for the reverberation time, T_{target} as a function of room usage and volume

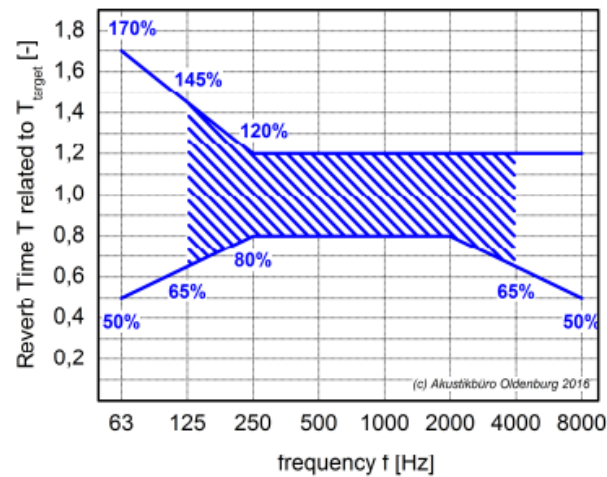
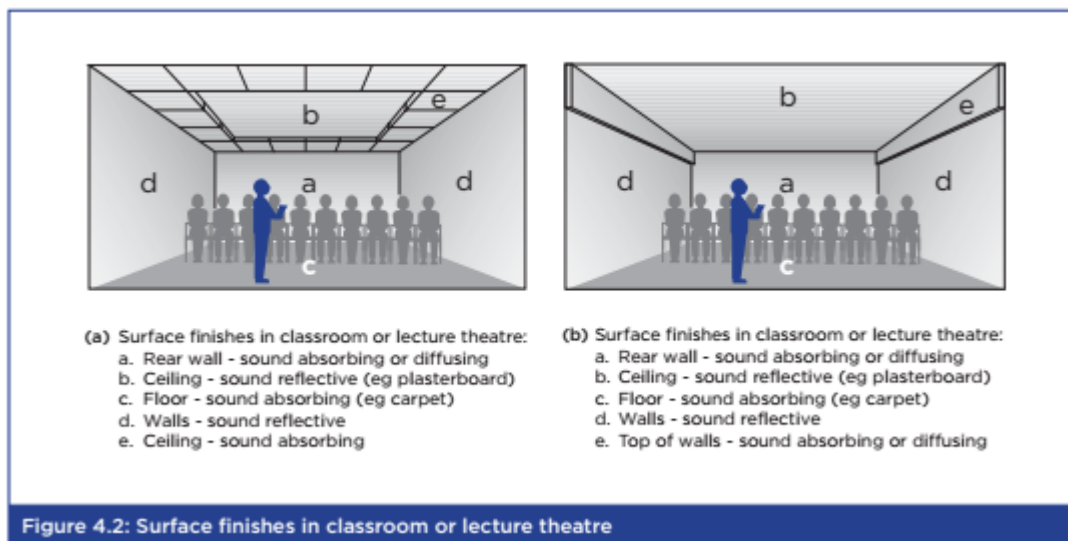
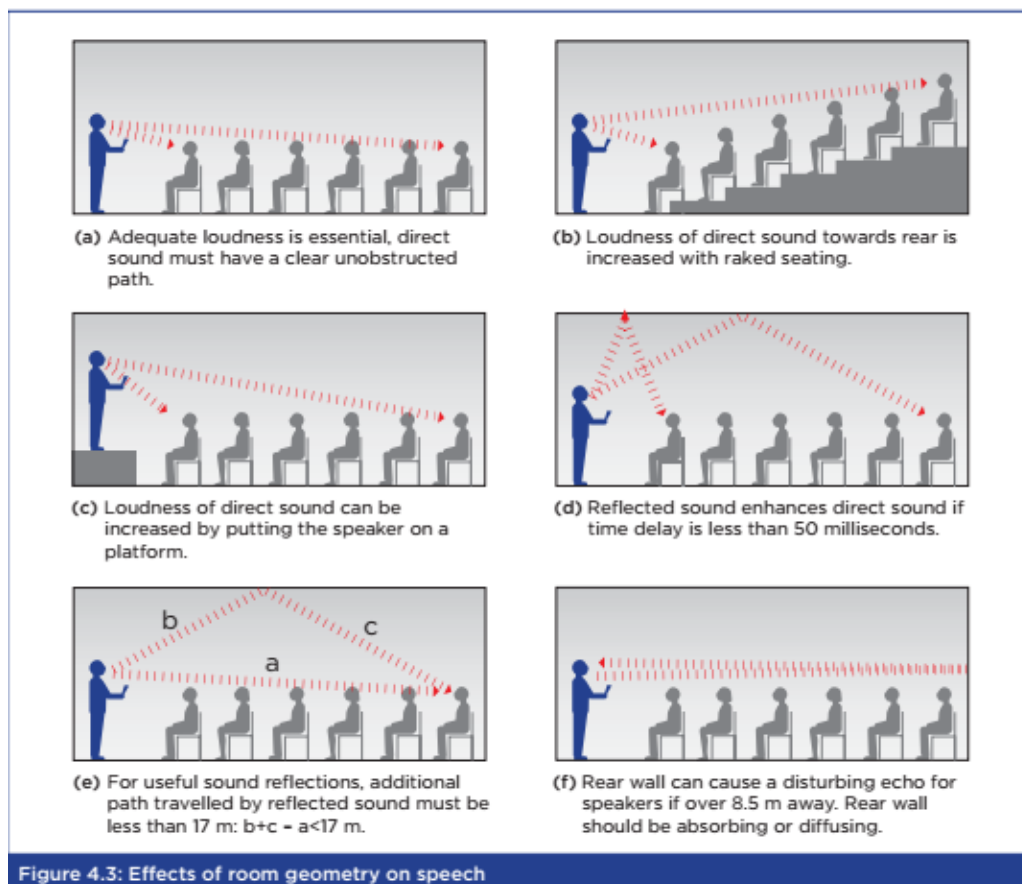


Figure 3: Tolerance range for the frequency dependent reverberation time in function of T_{target}

Аналогічною є схема розташування поглинаючих поверхонь, наведена в [6] (Figure 4.2).



В цьому ж джерелі наведено схему (Figure 4.3), яка обґрунтовує таке розташування.



Рівні шумів для шкільних приміщень запропоновано в роботі [5] (Table 1).

Table 1: noise activity and sensitivity levels and upper limits for indoor ambient noise level

Type of room	Room classification for the purpose of airborne sound insulation in Tables 3a and 3b		Upper limit for the indoor ambient noise level $L_{Aeq,30mins}$ dB	
	Activity noise (Source room)	Noise tolerance (Receiving room)	New build	Refurbishment
Nursery school rooms <i>Primary school:</i> classroom, class base, general teaching area, small group room <i>Secondary school:</i> classroom, general teaching area, seminar room, tutorial room, language laboratory	Average	Medium	35	40

Література до лекції 2.1

[1] Nocke, C. (2016). New standards in Architectural Acoustics - a German view. ACOUSTICS`2016, Brisbane, Australia. Available in: https://www.acoustics.asn.au/conference_proceedings/AASNZ2016/papers/p114.pdf

[2] DIN 18041:2004-05. (2004). Hörsamkeit von kleinen und mittleren Räumen. (Acoustical quality in small to medium-sized rooms.),” Beuth Verlag GmbH. Available in: <https://www.beuth.de/de/norm/din-18041/69640391>

[3] ANSI/ASA S12.60. (2015). Acoustical performance criteria, design requirements, and guidelines for schools. Part 1: permanent schools. American National Standards Institute, Inc. (ANSI) and Acoustical Society of America (ASA) New York, USA. Available in: <https://webstore.ansi.org/standards/asa/ansiasas12602010partr2015>

[4] International standard ISO 3382-1. (2009). Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces. First edition. Available at: <https://www.iso.org/obp/ui/#iso:std:iso:3382:-1:ed-1:v1:en>

[5] British standard BS EN 60268-16:2011. (2012). Sound systems equipment - Part 16: Objective rating of speech intelligibility by speech transmission index. BSI. Available at: https://infostore.saiglobal.com/en-us/Standards/BS-EN-60268-16-2011-217511_SAIG_BSI_BSI_512663/

[6] Acoustic design of schools: performance standards. (2015). Building bulletin 93. UK Department for Education, UK Education Funding Agency. Available at: <https://www.bloconconsulting.co.uk/building-bulletin-93-bb93/>

[7] Canning, D., James, A. (2012). The Essex Study. Optimised classroom acoustics for all. The Association of Noise Consultants. Available at: https://www.ndcs.org.uk/media/1808/the_essex_study_optimised_classroom_acoustics_for_all1.pdf

[8] IEC 60268-16. (2016). Sound system equipment – Part 16: objective rating of speech intelligibility by speech transmission index. International Electrotechnical Commission (IEC), Geneva, Switzerland. Available at: https://infostore.saiglobal.com/en-us/Standards/IEC-60268-16-4-0-569917_SAIG_IEC_IEC_1301856/

[9] Waterhouse, R. (1955). Interference Patterns in Reverberant Sound Fields. J. Acoust. Soc. Am., 27 (2). doi: <https://doi.org/10.1121/1.1939264>

Лекція 2.2. Методи активної корекції акустики приміщень

Нижче наведено огляд методів активної корекції акустики приміщень, виконаний за роботою *Stefania Cecchi, Alberto Carini and Sascha Spors. Room Response Equalization—A Review. Appl. Sci. 2018, 8, 16; doi:10.3390/app8010016. - P.1-47.*

Функція вирівнювання відгуку приміщення спрямована на поліпшення відтворення звуку в приміщеннях шляхом застосування технологій цифрової обробки сигналів з розробки еквалайзера на основі одного або декількох вимірювань відгуку приміщення. Ця тема інтенсивно вивчалася останні 40 років, в результаті чого був розроблений ряд ефективних методів, які вирішують різні аспекти цієї проблеми. Дана оглядова робота спрямована на те, щоб розглянути існуючі методи, а також обговорити їх плюси і мінуси. Огляд завершується обговоренням виникаючих тем і тенденцій.

Вступ

Коли звук відтворюється одним або декількома гучномовцями, сприйняття звуку змінюється в залежності від приміщення. В деякій мірі це можна розглядати як позитивний момент, оскільки додаються об'ємність і глибина, але довкілля та система відтворення звуку також можуть привносити небажані артефакти. Надмірні відбиття або резонанси у середовищі прослуховування можуть привести до небажаної зміни сприйняття звуку. Неідеальна система відтворення може навіть додати деякі артефакти (наприклад, зміна частот, нелінійність) до вихідного звуку.

Функція вирівнювання відгуку в приміщенні (RRE) вивчалася в теорії і застосовувалася на практиці для поліпшення якості відтворення звуку, протиставляючи згубний вплив навколишнього середовища в приміщенні і системи відтворення. В системі RRE функція передачі звуку в приміщенні (RTF), що характеризує шлях від системи відтворення звуку до слухача, вирівнюється за допомогою еквалайзера відповідної конструкції, який може бути реалізований декількома способами. Основна ідея полягає в тому, щоб

виміряти імпульсну характеристику приміщення (RIR) за допомогою мікрофона, а потім отримати еквалайзер з допомогою його інверсії. Однак на цей метод впливає декілька чинників, і тому за останні 40 років було розроблено велику кількість різних методів. Маємо знати, що в літературі для RRE використовувалося багато різних назв, таких як «вирівнювання приміщення», «корекція приміщення», «компенсація приміщення», «інверсія приміщення», «дереверберація приміщення», «дереверберація», «зменшення реверберації» і інші. У цьому огляді збірний термін RRE використовується для позначення будь-якої техніки, спрямованої на створення еквалайзера на основі вимірювань RTF.

Запозичуючи слова з [1], можна сказати, що існує «багатовимірність альтернатив для дизайну зворотного фільтра приміщення». Зокрема, інверсія RIR може виконуватися з урахуванням непараметричного підходу, такого як метод найменших середніх квадратів або пряма інверсія частотної характеристики [1,2], параметричного підходу, такого як моделювання авторегресивно-ковзаючого середнього (ARMA) моделювання [1,3], або контроль часового розпаду на низькій частоті [1,4]. Однак, як описано в [1], це не єдина можлива класифікація: RRE також можна розділити на методи мінімальної або змішаної фази. Перший спрямований тільки на вирівнювання величини RTF, в той час як інший також діє на надлишкову фазу RTF.

У цій оглядовій роботі представлена загальна класифікація, спрямована на більш широкий погляд на сучасний стан RRE. На рис.1 представлена концептуальна схема цієї класифікації, яка об'єднує різні методи, які будуть представлені нижче. Як показано на рис.1, підходи RRE діляться на одноточкові (один вхід/один вихід – SISO, кілька входів/один вихід – MISO) і багатоточкові (один вхід/кілька виходів – SIMO, еквалайзери для приміщень з кількома входами/виходами – MIMO). Однопозиційний кімнатний еквалайзер оцінює вирівнюючий фільтр на основі вимірювання в одному місці RTF [5]. Він ефективний тільки в обмеженій зоні навколо вимірюваної точки (розміром з частину довжини акустичної хвилі). Насправді RTF значно варіюється в

залежності від положення в кімнаті [6,7] і часу [2], оскільки кімнату можна розглядати як «слабку нестационарну» систему [8]. Для збільшення зони вирівнювання і для контрасту з варіаціями реакції приміщення були запропоновані багатоточкові еквалайзери [9]. Багатоточковий еквалайзер для кімнати використовує декілька вимірювань RTF в різних місцях для створення еквалайзера. Ці підходи можна використовувати для фіксованої та адаптивної корекції. Перший підхід, що ґрунтується на RTF, вимірюється в фіксованих положеннях в певний час. Останній здатен відстежувати і пристосовуватися до змін у відгуку приміщення через його часову характеристику, що змінюється, та що виникає, наприклад, через зміни температури або руху людей або інших перешкод. Різні методи попередньої обробки застосовуються для контрастування чутних спотворень, викликаних фіксованою корекцією в сценаріях, де RTF змінюються. Також можуть бути застосовані різні методи проектування еквалайзера (далі класифікуються як мінімально-фазові або змішані). Зовсім недавно [10] була досліджена еквалізація в просторово-часових перетворених областях для адаптивної еквалізації масивних багатоканальних систем відтворення звуку, і в даний час ця тема активно досліджується.

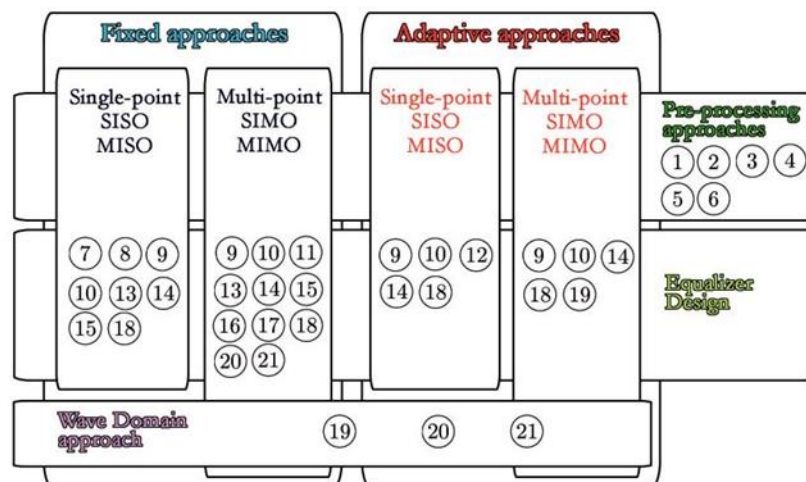


Рис. 1. Загальна класифікація систем вирівнювання реакції приміщення (RRE). Можливі підходи: 1 – короткі фільтри, 2 – комплексне згладжування, 3 – перетворення частоти, 4 – фільтри Каутца, 5 – різношвидкісні підходи, 6 – перетворення імпульсної характеристики приміщення (RIR), 7 – гомоморфні фільтри, 8 – лінійний прогнозний аналіз коду, 9 – оптимізація за методом найменших квадратів, 10 – методи деконволюції в частотній області, 11 – рішення зворотної теореми з множинними входами і множинними виходами (MINT), 12 – методи середнього і зваженого

середнього, 13 – методи кластеризації, 14 – методи прототипу, 15 – загальна акустична компенсація полюсів, 16 – модальне вирівнювання, 17 – підхід плоских хвиль, 18 – квазібезеховий підхід, 19 – адаптивні фільтри в хвильовій області, 20 – підхід до області перетворення, 21 – метод з урахуванням геометрії приміщення; MIMO: кілька входів/виходів; MISO: кілька входів/один вихід; SIMO: один вхід/кілька виходів; SISO: один вхід/один вихід.

Ця стаття спрямована на надання сучасного огляду RRE, обговорення плюсів і мінусів кожного методу з урахуванням історичної еволюції. Варто підкреслити, що проблема RRE аналізується з точки зору аналізу імпульсних характеристик. Всі підходи, які не засновані безпосередньо на аналізі RIR (наприклад, параметричні або графічні еквайзери), не обговорюються. Читача відсилаємо на [11] для всебічного огляду по цій темі. Ще одна область досліджень, пов'язана з RRE, яка не розглядається в цій статті, – це просторове звучання. Читач відсилається до [12], там є недавній огляд.

Ця оглядова стаття організована таким чином: Розділ 2 описує характеристики імпульсних відгуків приміщення і їх сприйняття слуховою системою людини. Розділ 3 вводить базову концепцію RRE, пояснюючи основні проблеми при інвертуванні реакції кімнати. У розділі 4 описуються підходи, використані для проектування еквайзера після їх історичного розвитку. У розділі 5 обговорюються методи попередньої обробки, що використовуються для усунення варіацій RIR з урахуванням людського сприйняття. Розділ 6 описує еволюцію від одноточкової до багатоточкової еквалізації з використанням декількох мікрофонів, розміщених в кімнаті. Розділ 7 описує адаптивні підходи до RRE в рамках одноточкової і багатоточкової еквалізації. Розділ 8 знайомить з новаторськими підходами після теоретико/хвильового погляду на проблему. Розділ 9 описує інструментальні і перцепційні заходи, використовувані для сучасної оцінки підходів RRE. Розділ 10 повідомляє про нові методи і нові тенденції в цій області. Нарешті, розділ 11 завершує цю статтю.

1. Реакція приміщення і його сприйняття

Характеристики реакції приміщення в часовій і частотній областях пов'язані з акустичними властивостями довкілля, які впливають на сприйняття людиною. У зв'язку з цим аспектом має сенс сформулювати аналіз імпульсної характеристики для вирішення важливих питань, які слід враховувати в процедурі RRE для досягнення поліпшення якості прослуховування звуку. Це включає в себе знання про людське сприйняття і психоакустики, які будуть явно використані в процедурі вирівнювання. Імпульсний відгук, отриманий від джерела звуку в певному місці реального навколишнього середовища, можна розділити на три частини [13]: (I) прямий звук; (II) ранні відбиття і (III) пізні відбиття, як показано на малюнку 2a. Перехід від ранніх відбиттів до пізніх відбиттів задається часом змішування, який оцінює час, що пройшов від ранніх відбиттів до пізніх. Його можна оцінити кількома способами [14,15]. Прямий звук і ранні відбиття мають фундаментальне значення для локалізації джерела звуку і сприйняття його тембру [16-18], в той час як пізня реверберація дає уявлення про розмір кімнати [19]. Дослідження по сприйняттю відбиттів і їх впливу на тембр можна знайти в [19-25]. Спектральний склад прямого і відбитого звуку різниться. Стіни, порт'єри і оббивка зазвичай поглинають високі частоти відбиттів. Ефект посилюється за рахунок множинних відбиттів, при цьому пізня реверберація зазвичай має набагато меншу енергію на високих частотах. На низьких частотах довжина хвилі порівняна з типовими розмірами приміщення: стоячі хвилі можуть з'являтися в приміщенні для стійких сигналів, що призводить до добре визначених залежних від положення максимумів і мінімумів амплітудної характеристики. На цих частотах відгук приміщення має плавну поведінку, що характеризується добре розділеними резонансами і виїмками, як показано на рис. 2b. Резонанси і виїмки визначаються інтерференційними картинами, викликаними прямим звуком і відбиттями, з виїмками, що з'являються, коли різниця в довжині шляху становить непарне число півхвиль. Насічки зі збільшенням частоти стають все більш щільними. Для частот, що перевищують частоту Шредера [13], частотна характеристика стає надзвичайно нерегулярною. Спектральні піки більш чутні, ніж виїмки [20],

але також чутні широкосмугові виїмки [26]. На високих частотах піки і виїмки сильно залежать від положення в кімнаті і від таких факторів, як вологість і температура в приміщенні [27-29] або руху перешкод [30-34]. Слід зазначити, що настільки великі відмінності у відгуку мало впливають на суб'єктивне враження слухача [18]. Було висловлено припущення [19,22,24], що вухо більш чутливе до початку сигналу (тобто до всього спектру початкової частини RIR) і що воно в значній мірі ігнорує високочастотні компоненти пізньої реверберації [35]. Цей аспект слід враховувати при розробці еквайзера.

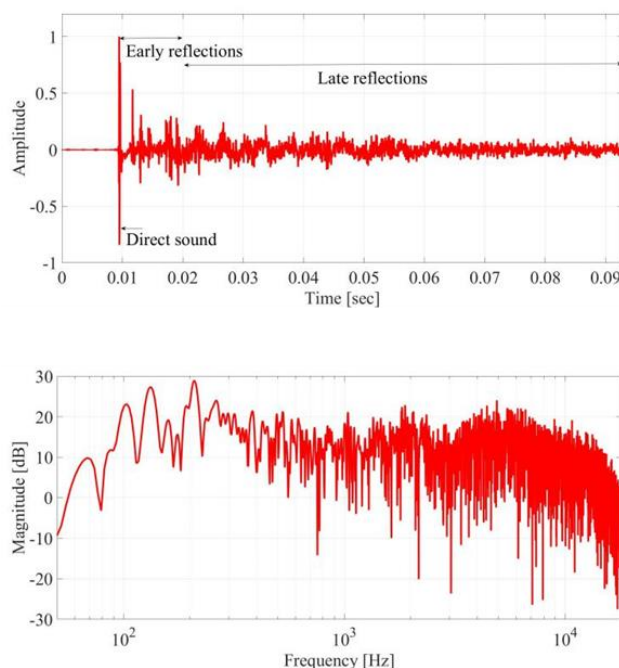


Рис. 2. Реальна поведінка RIR (а) в часовій області і (б) в частотній області

На сприйняття високих частот особливо впливає частотна роздільна здатність слухової системи людини. Роздільна здатність вуха є нелінійною і неоднорідною по частоті, з майже логарифмічною залежністю від частоти [36]. Цей аспект призвів до введення психоакустичних частотних шкал в розробку еквайзера з метою зміни спектрального змісту у відповідності до людського сприйняття. Шкала Мела [37], шкала Барка (критична смуга пропускання) [38] і шкала ERB (еквівалентна прямокутна смуга пропускання) [39] є прикладами психоакустичних частотних шкал, які зазвичай ґрунтуються на моделі слуху з набором фільтрів. Шкала Мела – це шкала сприйняття піків, яка оцінюється слухачами, які знаходяться на рівній відстані один від одного. Шкала Барка

заснована на критичних смугах. ERB також пов'язана зі шкалою Барка і слуховими фільтрами, оскільки фільтри ERB пропускають таку саму кількість енергії, що і слухові фільтри, яким вони відповідають. Можна зробити висновок, що логарифмічна шкала частот людського слуху багато в чому пояснює низьку чутливість до піків і провалів на високих частотах, і цей аспект слід враховувати при розробці еквалайзера.

Часова інтеграція і маскуючі властивості слухової системи людини також впливають на сприйняття відбиттів. Вухо сприймає звуки, об'єднуючи їх з «вікном» тривалістю близько 60 мс, яке має еквівалентну прямокутну тривалість 5 мс [40]. «Вікно» асиметричне, з більш повільним підйомом і більш швидким згасанням. Вухо нечутливе до часових подій тривалістю менше 2 мс [41]. Маскування вказує на стан, при якому звуки, які представлені ізольованими, будуть чутні, приховані присутністю звуку більш високого рівня (маскуючого). У нас може бути як одночасне, так і неодночасне маскування. Одночасне маскування залежить від частоти маскуючого і маскувального сигналу. Максимальний ефект досягається, коли вони відрізняються менш ніж на критичну смугу пропускання. Ефект швидко зменшується, коли частота маскуючого сигналу більше, ніж замаскований сигнал, і зменшується повільніше, коли частота маскуючого сигналу нижче, ніж замаскований сигнал [42,43]. Неодночасне маскування відноситься до ситуацій, коли маскуючий звук і маскувальний звук розділені в часі.

Він поділяється на зворотне маскування із замаскованим звуком, що передує маскувальному, та пряме маскування, тобто маскуючий звук слідує за маскувальником. Обернене маскування є досить обмеженим у часі [43]: його ефект зникає через 15-20 мс [44,45], а найбільш значна частина згасає через 5 мс [46]. Попереднє маскування має довше розширення на 100-200 мс. Його поведінка схожа на одночасне маскування, і це залежить від частотного співвідношення між маскувальним та маскованим звуком. Згідно з [39], його ефект починається як одночасне маскування, а потім з часом зникає прямолінійною лінією на графіку, що представляє зменшення маскування в дБ

проти часу [43]. Середня крива маскування вперед була введена в [43,47]. У перші 10 мс крива має постійне значення, рівне —9 дБ, що є максимальним рівнем маскування в [19], а потім вона з часом спадає. Це явище можна використати в процедурі вирівнювання, як це обговорюється в наступних розділах.

Чутність кімнатних відбиттів також залежить від напрямку надходження прямого звуку та віддзеркалень щодо слухача [48], від гучності прямого звуку (відбиття можна легше сприймати при голоснішому прямому звуці), від виду сигналу [19], а також на спектральний вміст прямого звуку та відбиття (маскування має сильніший ефект, якщо спектральний вміст прямого звуку та відбиттів збігається) [48].

У наступних розділах обговорюються різні методи RRE, що висвітлюють проблеми, що впливають із фізичних властивостей реакції приміщення та як можна включити характеристики слухової системи людини.

3. Зворотність відгуку кімнати

Першу дослідницьку роботу з RRE можна віднести Нілі та Аллену в 1979 р. [5]. У своїй фундаментальній роботі вони вивчали зворотність RTF та її наслідки. Розглядаючи RTF синтетичного приміщення, вони показали, що якщо відбивна здатність стіни низька (нижче 36%), RTF є мінімально-фазовою і, таким чином, зворотною. Навпаки, при більших коефіцієнтах відбиття стін, як у типових приміщень (в діапазоні 70% -90%), RIR не є мінімальною фазою. Однак усе ще можливо вирівняти мінімально-фазову частину відгуку приміщення (тобто амплітудну характеристику та мінімально-фазову частину фазової характеристики), врахувавши RTF $H(z)$ у добуток мінімально-фазної термін $H_m(z)$ і стабільний всепрохідний фільтр $A(z)$,

$$H(z) = H_m(z) A(z). \quad (1)$$

Фільтр еквалізування просто обчислюється шляхом прийняття зворотного z -перетворення, зворотного до спектру $H_m(z)$. Слухаючи результат вирівнювання мінімальної фази, Нілі та Аллен повідомили, що "Ефект кімнати

був видалений, але тон, схожий на дзвінок дзвінка, лунав у фоновому режимі" [5].

На оригінальний підхід [5] насправді впливає кілька проблем, багато з яких виявлені дослідниками лише в подальших дослідженнях. Дотримуючись хронологічного порядку, в якому вирішувались ці проблеми:

- Коли реакція приміщення не є мінімально-фазовою, точне зворотне не може бути реалізоване з одним джерелом звуку, оскільки інверсія є або нестійкою, або акаузальною.
- Точне вирівнювання відгуку приміщення або його мінімально-фазової частини вимагає дуже довгих фільтрів.
- На еквалайзер впливає будь-яка недосконалість вимірювання реакції приміщення [6,34].
- Реакція приміщення сильно залежить від розташування гучномовця та мікрофона, що використовується для вимірювання [6,31-34,48,49].
- Точне вирівнювання можливо лише в одному місці, а протяжність зрівняльної зони - це лише частка довжини акустичної хвилі [6]. На високих частотах зона вирівнювання може бути меншою, ніж міжвушна відстань вух (близько 18 см).
- Вирізи реакції приміщення, на які впливає рівень шуму, сильно підсилюється еквалайзером із генерацією часто чутного звуку, схожого на тон (дзвінок дзвінка, який відчувають Нілі та Аллен) [50-52].
- Реакція приміщення насправді повільно змінюється в часі, на неї впливають вологість і температура [28,29], а також переміщення людей або інші перешкоди в огорожі.
- Людське вухо чутливе до надлишкової фази RTF [53].
- Еквалайзер повинен зберігати природне заспокоєння гучномовців на низьких і високих частотах [54,55].

Посилення цих частот може спричинити неприродне посилення реакції гучномовця, спричиняючи нелінійні ефекти, розсіювання енергії та можливі пошкодження.

Далі ми обговоримо різні рішення, які були розроблені для протиставлення вищезазначеним проблемам. Зокрема, ми розглянемо методи, що використовуються для проектування еквалайзера, враховуючи як стратегію вирівнювання мінімальної фази, так і змішану фазу, а також методи попередньої обробки, що використовуються для протиставлення ефектів варіацій реакції приміщення з положенням та часом. Наскільки це можливо, ми спробуємо слідувати хронологічному порядку, в якому були запропоновані методики, щоб проілюструвати еволюцію PPE.

4. Техніка проектування еквалайзера

В обговорюваних нами техніках еквалайзер кімнатної реакції розроблений на основі вимірювань RIR або RTF в одному або декількох місцях у бажаній зоні прослуховування. Як ми побачимо в Розділі 5, реакція приміщення в більшості випадків попередньо обробляється для протиставлення деяких шкідливих наслідків, про які йдеться у Розділі 3. У будь-якому випадку, прототип відгуку приміщення зазвичай отримують і використовують для дизайну еквалайзера.

Більшість методів проектування еквалайзера можна класифікувати на наступні п'ять класів:

- Гомоморфна фільтрація;
- Аналіз лінійного прогнозного кодування (LPC);
- методи оптимізації найменших квадратів (або інші);
- деконволюція частотної області;
- Рішення з оберненою теоремою (MINT) із кількома входами / з кількома виходами.

Перші дві методики зазвичай використовують для мінімального фазового вирівнювання, останні три для змішаного фазового вирівнювання.

4.1. Гомоморфна фільтрація

Гомоморфна фільтрація вже була запропонована для мінімального вирівнювання фаз у початковій роботі Нілі та Аллена [5], але багато інших авторів представили модифіковані версії гомоморфної техніки [56-58]. При гомоморфній фільтрації мінімальна фазова частина реакції приміщення

виглядає з причинно-наслідкової частини комплексної цепструми. Потім отримують стабільний еквалайзер нескінченної імпульсної характеристики (IIR) шляхом прямої інверсії частини мінімальної фази. Оскільки було виявлено, що частина надлишкової фази RTF несе більшу частину енергії, що відбивається [59], в [6,56] гомоморфна техніка також застосовувалася для змішування фаз. Зокрема, мінімальний фазовий еквалайзер був доповнений еквалайзером надлишкової фази, розробленим за методикою найменших квадратів. Іншою можливістю реалізації еквалайзера надлишкової фази є використання відповідного фільтра, тобто фільтра, що має імпульсну характеристику, яка є зворотом часу імпульсної характеристики в системі надлишкової фази [57]. Однак змішане фазове вирівнювання, засноване на гомоморфній техніці, виявилось надмірно чутливим до помилок у початковому гомоморфному розкладанні реакції приміщення [56,60]. Про вдосконалення гомоморфної техніки повідомлялося в [57] та [58]. У роботі [57] запропоновано ітераційний гомоморфний метод шляхом ітераційного згладжування величини RTF-характеристики. Методика долає потенційні числові проблеми і "забезпечує більше розуміння суб'єктивних аспектів вирівнювання величини та фази при зменшенні акустичної реверберації" [57]. У роботі [58] перед обчисленням зворотного фільтра деякі з домінуючих полюсів низької частоти функції передачі фільтра замінюються новими з меншою величиною. Методика дозволяє зменшити ступінь коливань, пов'язаних з цими полюсами. Основним недоліком гомоморфної техніки є велика довжина абсолютно нульової (кінцевої імпульсної характеристики) моделі відгуку приміщення та висока чутливість моделі до "змін у розміщенні джерела / приймача" [61]. З цієї точки зору, аналіз LPC забезпечує більш надійні результати [61].

4.2. Аналіз LPC

При аналізі LPC відгук приміщення моделюється мінімально-фазним всеполюсним фільтром, а еквалайзер - фільтром кінцевої імпульсної характеристики (FIR). Модель усіх полюсів може бути отримана різними методами, включаючи ефективний алгоритм Левісона-Дурбіна [62]. LPC-аналіз

був одним із найуспішніших підходів до мінімального фазового вирівнювання і був успішно використаний багатьма дослідниками [61,63-75]. Всеполюсний фільтр може адекватно моделювати спектральні піки відгуку приміщення, тоді як забезпечує менш точну модель вирізів. Слід пам'ятати, що в слуховій системі людини спектральні піки чутніші, ніж виїмки [20]. Більше того, реакція приміщення суттєво варіюється щодо положення відповідно до вирізів [49]. Поліусний еквалайзер може компенсувати найбільш чутні частини відгуку приміщення (спектральні піки), не збільшуючи виїмки, що є ще однією бажаною властивістю еквалайзера.

Основним обмеженням аналізу LPC є той факт, що його можна використовувати лише для вирівнювання мінімальної фази, і його потрібно доповнити іншими методами для вирівнювання надлишкової фази.

4.3. Методи оптимізації найменших квадратів

Змішане фазове вирівнювання вимагає наближення оберненого до не мінімального фазового відгуку, який є акаузальним. З метою наближення акаузальної імпульсної реакції це було запропоновано в [76] додати затримку у відповідь еквалайзера та спроектувати еквалайзер шляхом мінімізації критерію похибки найменших квадратів. Далі підхід, запропонований у [76], дотримувався і вдосконалювався багатьма дослідниками як для однопозиційного, так і для багатопозиційного вирівнювання [77] (див. Розділ 6). Змішане фазове вирівнювання вимагає введення затримки еквалайзера. Ця затримка повинна бути якомога меншою (близько декількох мілісекунд відповідно до характеристик маскування назад вуха [46]), оскільки це може спричинити дратівливі артефакти у вигляді попереднього дзвінка або попереднього відлуння ефекти. У той же час затримка повинна бути достатньо великою, щоб отримати розумне змішане фазове вирівнювання. Оптимізація найменших квадратів була ключовим елементом багатьох адаптивних рішень, починаючи з основної роботи [77], як це докладно описано в Розділі 7. Інші критерії оптимізації найменших квадратів, що враховують подальші обмеження, також були запропоновані; наприклад, деконволюція з регуляризацією [51],

переформатування відгуку приміщення [78], фільтри Кауца [55] та короткі фільтри [79].

Основними обмеженнями методів найменших квадратів є висока чутливість до піків та виїмок відгуку приміщення, нерівномірний розподіл помилок у спектрі та можливість артефактів попереднього дзвінка або попереднього відлуння, спричинених еквалайзером затримка.

4.4. Деконволюція частотного домену

Інша техніка, яка використовується для проектування еквалайзера, заснована на деконволюції частотної області. Як спочатку пропонувалося в [80], еквалайзер може бути безпосередньо розроблений в дискретному перетворенні Фур'є (DFT), враховуючи зворотну відповідь кімнати. У роботі [80] метод застосовується до DFT віконної імпульсної характеристики, щоб виправити лише ранні відображення відгуку кімнати, що впливають на сприйняття тембру, і отримати короткий відгук еквалайзера. Загалом, цю техніку можна застосовувати як для вирівнювання мінімальної фази, так і для змішаної фази (додаючи відповідну затримку), але реакція приміщення повинна бути попередньо оброблена. Зокрема, глибину вирізів відгуку приміщення слід належним чином обмежити, щоб уникнути надмірних посилень та тривалих імпульсних реакцій еквалайзера, що може призвести до тональних артефактів [52]. У роботі [50] еквалайзер розроблений шляхом ділення комплексного спектра цільової реакції на складний спектр вимірюваної реакції приміщення. Щоб уникнути проблеми з надрізами, до вимірюваної реакції приміщення додається позитивний зсув. Цей прийом відомий як "деконволюція з регуляризацією", з ухилом, що називається "параметром регуляризації". Концепція була формалізована Кіркебі та його колегами в [51,81], розглядаючи критерій оптимізації найменших квадратів із "штрафним зусиллям". Він також відомий як "алгоритм Кіркебі". У [52] методика була застосована до RRE. Параметр регуляризації на практиці контролює найдовшу постійну часу зворотних фільтрів [81]. Щоб забезпечити, щоб константа часу не була ні занадто довгою, ні занадто короткою, параметр регуляризації повинен бути

встановлений відповідним чином [51]. У роботі [82] автори показують, як на полюси розчину деконволюції впливає параметр регуляризації. Зокрема, для кожного нуля, близького до одиничного кола, генерується триплет із двох полюсів і одного нуля, причому один з полюсів знаходиться поза одиничним колом. Цей полюс відповідає за акаузальну реакцію, і тому слід ввести моделювання затримки. У роботі [43] наводиться аналіз RRE на основі алгоритму Кіркебі на основі психоакустичних критеріїв. В розглянутих умовах було показано, що "помилки в процесі деверберації виявились надзвичайно чутними та дратує резонансами. Вони виникли внаслідок наявності глибоких спектральних вирізів у передавальних функціях комбінацій гучномовців, що створило тональні артефакти, що мали місце задовго до і після звуків прямого прибуття. Крім того, була виявлена надзвичайна чутливість до змін положення, що перешкоджало оптимізації дереверберації на практично розмірних зонах прослуховування. Встановлено, що якість дереверберації ще більше погіршується для більших акустичних просторів. " Незважаючи на ці обмеження, деконволюція з підходами регуляризації успішно застосовується у поєднанні з іншими методами, що використовуються для уникнення відчутних спотворень. Наприклад, його поєднували з перекосом частоти [83,84] або використовували в синтезі хвильового поля [85].

4.5. Методи зворотної теореми із кількома входами / з кількома виходами

Метод точної інверсії RIR - навіть коли це не мінімальна фаза - був запропонований в [86,87]. Метод базується на принципі, який називається зворотною теоремою з кількома входами / з кількома виходами (MINT). За допомогою цього методу інверсія будується з декількох фільтрів FIR, додаючи "деякі додаткові канали передачі акустичного сигналу, що створюються кількома гучномовцями або мікрофонами". На практиці MINT заявляє, що можна отримати точну інверсію відгуку приміщення, якщо кількість гучномовців більше, ніж кількість мікрофонів (тобто точок вимірювання). Таким чином, підхід є внутрішньо багатоканальним. Розглянемо випадок системи з двома гучномовцями та одним мікрофоном. Вкажемо за допомогою $G1(z)$ та $G2$

(z) функції передачі від гучномовців до мікрофонів, а за допомогою $H_1(z)$ та $H_2(z)$ функції передачі еквалайзерів, пов'язаних з кожним гучномовцем. Тоді для точної інверсії відгуку кімнати $H_1(z)$ та $H_2(z)$ повинні задовольняти наступній умові:

$$H_1(z)G_1(z) + H_2(z)G_2(z) = 1.$$

Як показано в [87], рішення рівняння (2) існує, якщо $G_1(z)$ і $G_2(z)$ відносно прості (тобто не мають спільних нулів), а коли рішення існує, то порядки $H_1(z)$ та $H_2(z)$ нижче, ніж $G_2(z)$ та $G_1(z)$, відповідно. Підхід є дуже потужним, оскільки дозволяє подолати проблему невикористаності еквалайзера. Однак підхід MINT також демонструє сильні обмеження. У роботі [88] MINT аналізується в числовій перспективі, вивчаючи номер умови інвертованої матриці часової області. Показано, що номер умови матриці часової області пов'язаний із особливими значеннями матриці передачі, що обчислюються за частотою. Кількість умов зменшується, а числові показники покращуються із збільшенням кількості гучномовців. Однак число умов збільшується "із швидкістю приблизно 1 біт" (тобто приблизно 6 дБ) для кожного доданого мікрофона [88]. Більше того, аналіз техніки MINT також представлений в [89], обговорюючи умови, які повинні бути виконані для існування точної матриці зворотного фільтра. Крім того, [89] демонструє, що кількість гучномовців повинна перевищувати кількість мікрофонів у спосіб, який відповідає висновкам [87]. Більше того, виведена явна формула, що вказує кількість необхідних зворотних коефіцієнтів фільтра для існування точного зворотного. Стаття також досліджує просторову протяжність зон вирівнювання, утворених зворотним фільтруванням. Показано, що вирівнювана зона масштабується за розміром відповідно до довжини акустичної хвилі при найвищій частоті, що нас цікавить.

Низький ступінь вирівненої зони та чисельна чутливість до помилок у вимірюваних відгуках, як видається, є основними обмеженнями MINTI. Удосконалення методу було запропоновано в [90], де розглядається більше контрольних точок без збільшення кількості зворотних фільтрів. Інше вдосконалення обговорюється в [91], де ітераційний метод застосовується до

МІНТ з урахуванням оптимально зупиненого зваженого градієнта спряженого. Для поліпшення обчислювальної ефективності MINT в [92] був представлений передискретизований піддіапазонний підхід із децимацією.

4.6. Альтернативна класифікація еквайзерів

Як пояснювалося раніше, еквайзери можуть бути класифіковані кількома способами, а вищезазначені техніки проектування вже класифіковані на мінімальні або змішані фази. Інша цікава класифікація методів проектування еквайзера була подана в [55]. Відповідно до [55], методи проектування еквайзера можна класифікувати на «непрямі» та «прямі» методи. Як показано на малюнку 3, непрямі методи оцінюють модель реакції приміщення - можливо, оброблену - для отримання еквайзера шляхом інверсії моделі. Натомість прямі методи мінімізують помилку між вирівняною реакцією приміщення та цільовою відповіддю. З цієї точки зору, гомоморфна фільтрація, LPC-аналіз та деконволюція частотної області є непрямими методами, тоді як оптимізація найменших квадратів - прямим методом. Методи оберненої теореми з кількома входами / з кількома виходами можна класифікувати як прямі, так і непрямі методи, оскільки вони обчислюють еквайзер, враховуючи інверсію матриці кімнатних відповідей. Однак, згідно з рівнянням (2), їх також можна оцінити, мінімізуючи похибку щодо ідеального відгуку.

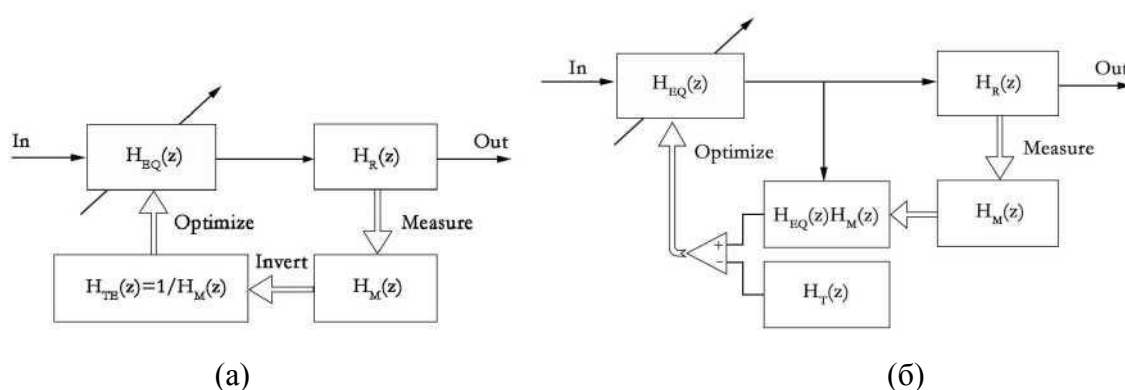


Рис. 3 (а) Непряма та пряма (б) класифікація методів проектування еквайзера, як повідомляється в [55], де HEQ являє собою фільтр вирівнювання, HR - канал відтворення, HM - виміряна імпульсна характеристика, а HR і HRE - цільові функції.

5. Методи попередньої обробки

Основні методи, розроблені для подолання обмежень RRE, що диктуються характеристиками реакції приміщення, а також з використанням психоакустичних властивостей слуху, розглядаються нижче. Ці підходи здатні модифікувати вимірюну RIR і повинні застосовуватися до фактичної процедури вирівнювання. Вони підходять як для одноточкового, так і для багатоточкового вирівнювання.

Основними проблемами RRE, які були вирішені на ранніх підходах, були дуже довгі імпульсні реакції еквалайзера, обмежена область простору, в якій RRE ефективна, і повільні коливання часу відповіді кімнати. Дуже довгий імпульсний відгук точного еквалайзера обумовлений спектральними характеристиками відгуку приміщення, як показано на малюнку 2b, з великою кількістю піків та вирізів, які збільшують їх щільність у напрямку до високих частот. Виїмки відповідають нулям поблизу одиничного кола в RTF. Таким чином, інверсний фільтр має полюси, близькі до одиничного кола, які визначають тривалий імпульсний відгук. Виїмки на високих частотах надзвичайно змінюються з положенням і часу, визначаючи малий обсяг в просторі та часі, в якому еквалайзер діє. Рухи положення прослуховування на 10 см можуть спричинити коливання реакції приміщення до 20 дБ [93], і для контрастування цих варіацій потрібна техніка попередньої обробки.

5.1. Короткі фільтри

Одним з перших доцільних для покращення RRE вдавсь до використання коротких фільтрів вирівнювання. Розглядаючи грубу модель реакції приміщення, яка намагається зафіксувати і виправити лише загальну тенденцію реакції кімнати, уникаючи моделювання різких піків та вирізів, можна зменшити тимчасову довжину імпульсної характеристики еквалайзера. Це рішення також корисно для збільшення масштабу зрівняльної зони та для подолання змін у часі відгуку кімнати [2]. Однією з найефективніших технік для проектування коротких фільтрів вирівнювання є те, що базується на аналізі LPC [61], який отримує хороше моделювання піків відгуку приміщення з більш грубим моделюванням насічок.

5.2 Нерівномірна роздільна здатність частоти

Для підвищення точності та ефективності вирівнювання еквайзер повинен використовувати характеристики реакції приміщення та людського вуха. На низьких частотах частотна характеристика кімнати є більш регулярною, а піки та виїмки в основному нечутливі до положення в кімнаті. Роздільна здатність вуха неоднорідна та нелінійна, з логарифмічною залежністю від частоти. На високих частотах вухо досить нечутливе до впадін характеристик кімнати і до високочастотної реверберації. Відповідно, еквайзер повинен забезпечувати точну роздільну здатність на низьких частотах і грубішу роздільну здатність на більш високих частотах. За цією стратегією розроблено багато методів:

- Комплексне згладжування,
- викривлення частоти,
- фільтри Kautz та паралельні фільтри IIR із фіксованими полюсами,
- Багатоступінчасті підходи.

5.2.1 Комплексне згладжування

Дробове згладжування октавного діапазону енергетичного спектра широко застосовується при обробці звуку. Його використання можна простежити з використанням аналогових еквайзерів (як, наприклад, аналізатори смуги фільтрів на третину октави), а пізніше було поширено на цифрові аналізатори спектру. У роботі [35,94] автори розширюють методику, вводячи методологію згладжування складної передавальної функції вимірюваної реакції приміщення дробовими октавними профілями. Методика може бути реалізована в часовій або частотній областях. Вона перцептивно сумісна, оскільки спектральне згладжування відбувається одразу за частотною роздільною здатністю вуха, з тонкою роздільною здатністю на низьких частотах і нижчою роздільною здатністю на високих частотах. Як результат, у часовій області застосування комплексного згладжування може зберегти початковий високочастотний вміст ранніх компонентів (тобто перехідну поведінку прямого звуку та перших відбиттів), а потім може поступово вводити низькочастотну фільтрацію

пізніших компонентів (тобто реверберації приміщення) [95]. Це також бажано з психоакустичної точки зору. У дисперсійному приміщенні вухо дуже чутливе до сигналів (тобто до повного діапазону частот першої частини RIR), тоді як воно менш чутливе до високочастотних компонентів пізнього відображення [19,22, 24,35]. Коли складні згладжені імпульсні відгуки використовуються для проектування RRE, вони дозволяють уникнути компенсації різких вирізів на високих частотах, щоб отримати зменшену довжину еквайзера, і вони забезпечують більш надійне вирівнювання з меншою чутливістю до можливих змін у позиції слухача та на інші варіації відгуку в кімнаті [2,35]. На малюнку 4 показано складний згладжувальний ефект на RTF для різних роздільних здатностей. Вводячи відповідну затримку еквайзера, комплексне згладжування дозволяє змішати фазове вирівнювання відгуку кімнати. Як альтернативу складному згладжуванню пропонується залежне від частоти сигналу віконне згладжування[96] або окреме згладжування величини та фази передавальної функції [97].

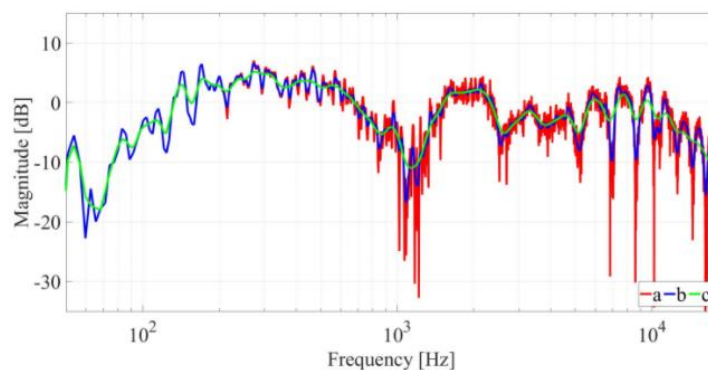


Рис.4 Комплексне згладжування вимірюваної функції передачі приміщення (RTF): (a) RTF; (b) комплексне згладжування на 1/12 октавної смуги; (c) комплексне згладжування на 1/3 октавної смуги.

5.2.2 Викривлення частоти

Іншим методом, що забезпечує неоднорідну роздільну здатність частоти, є “викривлення частоти” [98]. Початкова ідея деформації частоти представлена в [99], де введено неоднорідне перетворення Фур'є. Методика полягає у заміні одиничної затримки (z^{-1}) цифрових фільтрів на широкополосний фільтр першого порядку,

$$D_1(z) = \frac{Z^{-1} - \lambda}{1 - \lambda Z^{-1}}$$

таким чином отримуючи білінійне відображення одиничного кола на собі. Ефект викривлення можна регулювати, щоб наблизити спектральне зображення вуха [100]. У [101] наводяться аналітичні вирази, що наближають шкалу Барка та ERB. Вони дозволяють дуже добре наблизити шкалу Барка та менш точно наблизити шкалу ERB, завдяки необхідній вищій роздільній здатності, особливо на низьких частотах. Ефект перекошу частоти можна легко скасувати, знову замінивши одиничну затримку (z^{-1}) на широкосмуговий фільтр.

$$\tilde{D}_1(z) = \frac{Z^{-1} - \lambda}{1 - \lambda Z^{-1}}$$

На рисунку 5 наведено приклад впливу перекошу частоти на RTF для різних значень λ . Читач повинен відзначити розширення діапазону низьких частот та стиснення високих частот, отриманих при позитивних значеннях параметра викривлення λ .

Деформовані фільтри FIR та IIR можна отримати, замінивши відбіту лінію затримки ланцюжком фільтрів повного проходу першого порядку, але хоча реалізація деформованих фільтрів FIR відбувається миттєво [66], деформовані фільтри IIR вимагають відповідних конструкцій, щоб уникнути вільної без затримки петлі [102]. Викривлені ФІР-фільтри суворо пов'язані з фільтрами Лагера [103], єдина відмінність полягає в тому, що у фільтрі Лагера є додатковий попередній фільтр, розміщений перед ланцюгом усіх проходів [66]. Логарифмічна шкала частоти також може бути апроксимована, але в цьому випадку весь прохідний ланцюг повинен бути замінений набором фільтрів, сформованим широкосмуговими фільтрами.

Викривлення частот використовувалось у багатьох аудіо-програмах, починаючи від аналізу LPC [100], вирівнювання звуку [104,105], вирівнювання гучномовців [106–108] та фізичного моделювання гітарних тіл [109], до функції передачі голови (HRTF) фільтрація [66,109]. Читач звертається до [66] для огляду методів викривлення частоти та їх застосування. У контексті RRE багато науковців використовували викривлення частоти при мінімальному фазовому

вирівнюванні для покращення продуктивності еквайзера шляхом розширення роздільної здатності на низьких частотах та стиснення її на високих частотах. У якості психоакустично мотивованої шкали частоти - у більшості випадків використовується шкала Барка. Наприклад, у підході [70,71] прототип відгуку приміщення спочатку викривляється до приблизного масштабу шкали Барка. Потім в деформованій області отримують всеполюсну модель відгуку кімнати за допомогою аналізу LPC. Врешті-решт, мінімальний фазовий еквайзер виводиться у часовій області шляхом деформації зворотної до всеполюсної моделі за допомогою (4). Основним недоліком цього підходу є висока вартість обчислення операції деформації частоти. У роботі [73,75] перетворення частоти було ефективно здійснено шляхом нелінійної вибірки швидкого перетворення Фур'є з високою роздільною здатністю (БПФ) прототипу кімнатної реакції.

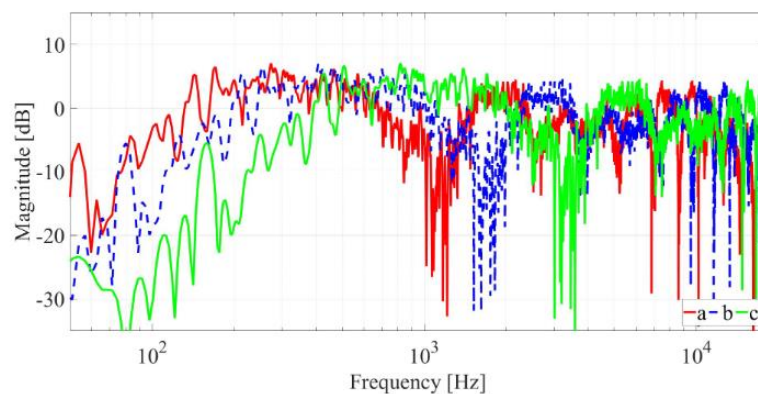


Рис 5. Частотна деформація вимірюваного RTF: (a) RTF; (b) деформований ПДК з $\lambda = 0,2$; (c) деформована RIR з $\lambda = 0,5$.

5.2.3 Фільтри Kautz та паралельні фільтри IIR із фіксованими полюсами

Фільтри Кауца - це раціональні ортонормальні структури фільтрів. Вони ортонормальні, оскільки мають ортонормальні імпульсні реакції. Функції раціональної передачі безперервного часу з ортонормальними імпульсними відгуками досліджував Каутц у [110]. Пізніше ортонормальні функції передачі дискретного часу були вивчені Брумом в [111], який назвав їх "дискретними функціями Кауца". Фільтри Кауца можна розглядати як узагальнення викривлених FIR-фільтрів та фільтрів Лаугерре, де ланцюг широкосмугових фільтрів з рівними полюсами замінений ланцюгом широкосмугових фільтрів з

окремими полюсами, можливо, складними [54]. На рисунку 6 показані результати моделювання Каутцом вимірюного RTF.

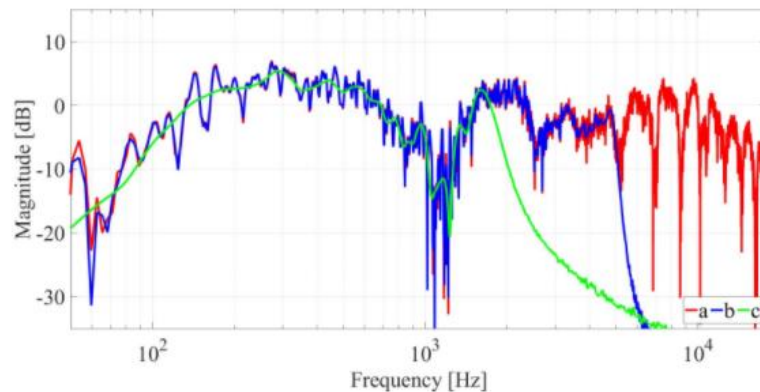


Рис.6 Викривлення частоти вимірюваного RTF: (a) RTF; (b) деформований RIR з $\lambda = 0,2$; (c) деформований RIR з $\lambda = 0,5$.

Правильно вибравши полюси, можна здійснити довільний розподіл частотного розподілу розробленого фільтра. Апроксимацію роздільної здатності шкали часових частот за допомогою фільтрів Кауца можна знайти в [55]. Полюси можна вибрати апріорі на основі бажаної роздільної здатності, але їх також можна налаштувати на конкретне застосування шляхом узгодження частоти полюсів з резонансами системи, що підлягає моделюванню [54]. На практиці тонка настройка полюсів необхідна при розробці моделей нижчого порядку для високорезонансних систем [54]. Після вибору полюсів система ідентифікована за допомогою фільтрів Кауца може отримати користь від ортогональності імпульсних реакцій. Читач звертається до [54] для обговорення методів встановлення полюсів та методів ідентифікації.

Фільтри Kautz використовувались для RRE, використовуючи неоднорідну частотну роздільну здатність цих фільтрів. Вони застосовувались як для мінімального, так і для змішаного фазового вирівнювання, використовуючи як нерухомі полюси, так і налаштовані полюси [54,55]. Коли застосовується фіксований полюсний підхід, фільтри Кауца також можуть бути розроблені та реалізовані у вигляді набору фільтрів секцій другого порядку [112–114], що має переваги для обчислювальної складності. У [114,115] показано теоретичну еквівалентність паралельних фільтрів та фільтрів Кауца, а також наведені

формули для перетворення параметрів двох структур одна в іншу. На рисунку 7 наведено приклад проектування паралельного фільтра, який відповідає методології [114].

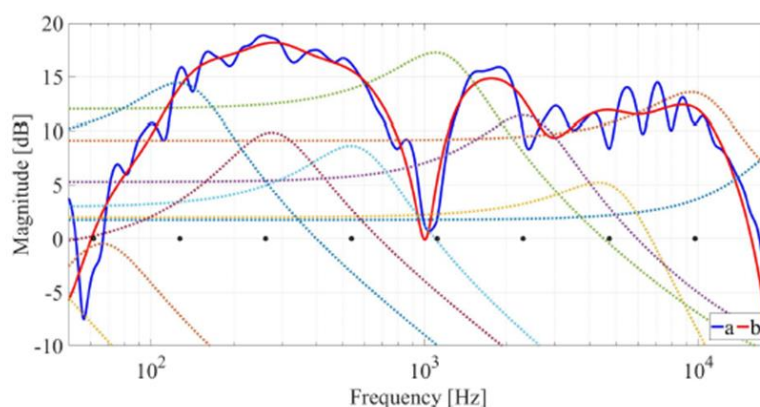


Рис 7. Приклад конструкції паралельного фільтра банку: (а) RTF; (б) результуюча частотна характеристика фільтра. Пунктирними лініями представлені окремі передавальні функції 16 секцій другого порядку, тоді як кола відображають полюсні частоти.

5.2.4. Багатоступінчасті підходи

Інша можливість досягнення неоднорідної роздільної здатності частоти дається багатоступеневими підходами. У цих підходах спектр ділиться на різні смуги, які беруть на вибірку та окремо обробляють фільтрами різної довжини. У більшості запропонованих підходів один із фільтрів охоплює низькі частоти [50,72,116–120], що використовується для модального вирівнювання та контролю низьких резонансів (див. Розділ 6.5) або для управління низькими частотами. Як правило, низькочастотні фільтри повинні компенсувати дуже тривалий час реверберації, і, таким чином, фільтри отримують вигоду від низької частоти дискретизації на низьких частотах. Фільтри, що використовуються для середніх і високих частот, зазвичай використовують нижчу роздільну здатність у порівнянні з низькими частотами, із значною економією обчислень.

Наприклад, у роботі [119] автори пропонують процедуру вирівнювання з двома смугами. Низькочастотний канал обмежується приблизно частотою Шредера через вибірку. Техніка на основі ШПФ з регуляризацією використовується для розробки мінімального фазового еквайзера з

гомоморфною фільтрацією. Верхня смуга також вирівнюється за допомогою міні-фазового еквалайзера, розробленого з використанням аналізу LPC та техніки викривлення. У [120] ті самі автори натомість розділили спектр на три смуги: низько-, середньо- та високочастотні. Низький діапазон знову обмежується приблизно до частоти Шредера - зокрема, 150 Гц - через дискретизацію вниз, але еквалайзер тепер розроблений за технологією LPC. У середньочастотному діапазоні від 150 Гц до 900 Гц еквалайзер розроблений за допомогою викривленої техніки LPC, щоб зосередити увагу на нижній частині смуги. Понад 900 Гц високочастотний спектр згладжується, щоб зменшити чутливість до положення, а потім еквалайзер знаходить шляхом інвертування згладженого спектра, накладаючи злегка зменшувану цільову функцію. Автори також поєднали цей базовий еквалайзер із еквалайзером надлишкової фази в низькочастотному діапазоні та попередньою обробкою, заснованою на техніці деконволюції в перші 10 мс після прямого звуку.

У роботі [72] автори поєднали техніку кластеризації багатоточкових нечітких с-середніх значень [121] (див. Розділ 6.2) з двосмуговим багатоступеневим підходом, відокремлюючи низькочастотну смугу нижче 80 Гц від високочастотної діапазон понад 80 Гц. Низькочастотна смуга зменшується в 256 разів для роботи з відгуками кімнати невеликої довжини до застосування методу кластеризації нечітких с-середніх для проектування еквалайзера.

5.3. Переформатування імпульсної реакції кімнати

Ще однією можливістю врахування психоакустичних характеристик вуха є переформування імпульсної реакції таким чином, що зміна приміщення стає нечутною. При скороченні RIR загасання вихідного RIR прискорюється, так що ефект реверберації послаблюється. У літературі пропонуються різні методи [78,122–130]. Далі ми розглянемо найбільш відповідні методи. Більшість із цих методів не є методами RRE у строгому розумінні, але їх можна легко поєднати з методами RRE. Перші спроби переробки RIR [122,123] намагалися адаптувати концепції скорочення каналів, розроблені в галузі телекомунікацій [131–134], застосовуючи алгоритми оптимізації найменших квадратів. Правильно

спроєктувавши переформований фільтр, можна максимізувати енергію вирівняного RIR у бажаному часовому вікні, мінімізуючи одночасно хвости реакції приміщення в небажаному вікні. Таким чином, наприклад, можна безпосередньо максимізувати міру D50 для розбірливості мови, яка є відношенням енергії протягом 50 мс після першого піку RIR до енергії повної реакції. Оптимізація найменших квадратів фільтра переформування, що відокремлює бажане часове вікно від небажаного вікна, дала незадовільні результати [122] у вигляді чутних пізніх відлунь або спектральних спотворень. Ці проблеми спричинені сильним розділенням, що накладається з урахуванням неперекриваючих бажаних та небажаних вікон, та оптимізацією найменших квадратів, що призводить до нерівномірного розподілу помилок. Таким чином, вже в [122] автори модифікували парадигму скорочення каналу з метою формування бажаної імпульсної реакції на коротший час реверберації, враховуючи поступовий перехід між бажаним і небажаним вікнами. Підхід було вдосконалено в [78,124,126,127]. Ці підходи використовують психоакустичні властивості слухової системи людини, зокрема ефект маскування вперед. Вони прагнуть отримати вирівняну реакцію, яка затухає досить швидко, щоб уникнути звукових ехо-сигналів, так що час реверберації маскується прямим звуком відповідно до ефекту маскування вперед слухової системи людини. Тут бажані та небажані вікна вказуються відповідно до середньої кривої маскування вперед [47] та [43]. Більше того, щоб уникнути проблем, пов'язаних з оптимізацією найменших квадратів, пропонуються оптимізація норми нескінченності та p -норми (з великими значеннями p). Цей підхід також застосовується до багатоканальних проблем у [125,130].

Жоден із згаданих підходів до перетворення RIR не вимагає спектральних вимог. У більшості випадків ці підходи зазвичай дають плоску загальну частотну характеристику, але при дуже довгих імпульсних відгуках вони можуть призвести до спектральних спотворень [128]. На противагу цій проблемі, в [128] цільова функція модифікована, щоб включити термін регуляризації на основі p -норми в частотну область, таким чином, накладаючи спільну оптимізацію в

часовій та частотній областях. У роботі [135] термін регуляризації замінено інтегрованою мірою спектральної рівності, що дозволяє інтегрувати концепцію слухових шкал у конструкцію еквайзера. Таким чином, підходи [128,135] поєднують переробку RIR із RRE.

6. Від одноточкового до багатоточкового вирівнювання

Інша класифікація RRE стосується кількості використовуваних мікрофонів або контрольних точок. Класичні підходи засновані на використанні одного RIR, захопленого поблизу позиції слухача (див. Розділ 4), маючи на увазі конкретну солодку пляму, де вирівнювання є ефективним [136]. Завданням багатоточкових еквайзерів є збільшення еквалізованої зони [137], також покращуючи стійкість еквайзера щодо похибок вимірювання та варіацій відгуку приміщення, неявно використовуючи варіації між кількома вимірами. Далі дається огляд методів вирівнювання за кількома крапками, беручи до уваги, що більшість методів, обговорених у попередніх розділах, також застосовувались до багатоточкових RRE.

6.1. Методи середнього та середнього зваженого

Один з найперших багатоточкових підходів був запропонований Елліотом та Нельсоном. У роботі [77] автори представили метод розробки фільтра вирівнювання для систем відтворення звуку шляхом регулювання коефіцієнтів фільтра таким чином, щоб мінімізувати суму квадратних помилок між вирівняними реакціями в декількох точках кімнати та відкладеною версією оригіналу сигналу. У статті розглядаються як фіксовані, так і адаптивні еквайзери на основі алгоритмів відфільтрованого x . Підхід є ефективним і застосовувався також у багатьох інших вдосконалених методиках [9,31,33,118]. Основне обмеження дається тим фактом, що неявне усереднення суми квадратних помилок не може використати подібність відповідей кімнати, а також не можна вирівнювати улюблене у певних положеннях. В контексті вирівнювання звуку в автомобілі [118], техніка була вдосконалена шляхом розгляду багатоточкового вирівнювання із середньозваженим значенням

помилки. Рішення забезпечило поліпшення реакції у вибраному місці, "без значних погіршень в інших точках" [118].

6.2. Методи кластеризації

Ми можемо використати подібність між різними просторово розподіленими відгуками кімнат, кластеризуючи їх відповідно до обраної міри відстані. У роботі [138] "надзвичайно великий набір" можливих RTF у корпусі був згрупований та зрівняний меншою кількістю еквалайзерів. RTF моделювались із всеполюсними фільтрами з використанням аналізу LPC, і таким чином були розроблені мініфазні еквалайзери. Потім було проведено векторне квантування для оптимальної класифікації всеполюсних фільтрів. Класифікацію можна використовувати як бібліотеку просторового вирівнювання, досягаючи зменшення частоти реверберації в широкому діапазоні позицій у приміщенні, залежно від фактичного положення слухача. Основним обмеженням цього методу є необхідність витягти та запам'ятати великий набір відповідей кімнати та еквалайзерів та відстежити позицію слухача.

Метод кластеризації нечітких с-середніх застосовується в [30,70–72,121,139,140]. У підході [121,139], "репрезентативні прототипові відповіді кімнати" отримані з кількох вимірних відповідей кімнати, які мають подібні характеристики, використовуючи метод нечіткого с-середнього неконтрольованого методу навчання. Потім прототипові відповіді об'єднуються, щоб сформувані "загальну точкову відповідь" на основі нечіткої стандартної адитивної моделі Коско [141, 142]. Метод використовує зважування відповідно до "рівня активації" прототипу, залежно від ступеня присвоєння відповідей приміщення кластеру, що містить прототип. Потім еквалайзер обчислюється з оберненої загальної точкової реакції за допомогою LPC-аналізу, «отримуючи значне поліпшення показників рівняння за просторовими методами усереднення» з придушенням піків у спектрах величини кімнати [139]. Далі метод був вдосконалений у роботі [70,71,140], застосувавши нечітку кластеризацію с-середніх значень до викривлених імпульсних реакцій, використовуючи, таким чином, перцептивні властивості вуха. Цей підхід також

був поєднаний з багатофункціональною фільтрацією в [72], щоб забезпечити ефективну фільтрацію низькочастотної характеристики при низьких частотах дискретизації з економією обчислень.

Пізніше підхід [70,139] був вдосконалений, застосувавши перекручування частот та нечітку кластеризацію середніх значень до реакцій кімнати за величиною [73,75], із значним покращенням з точки зору складності обчислень. Також зважена нечітка кластеризація середніх значень була запропонована в [143], де зразки RIR зважувались різним чином, щоб врахувати різний вплив, який вони мають на RRE.

Нечіткий підхід кластеризації s -означає [70, 139] також є першим прикладом «багатоточкового прототипу». Ці методи використовують вимірювання реакції приміщення в різних місцях для отримання прототипу відповіді, який є репрезентативним для сприйнятливої акустичної ситуації, яку потрібно виправити. Потім на основі цього відповіді прототипу розробляється одиничний еквалайзер непрямыми або прямими методами [55].

Різні підходи для визначення відповіді прототипу були вивчені в [144]. Зокрема, метод нечітких s -середніх порівнювали із середнім середнім значенням, медіаною, \min - \max та середньою квадратичною середньою величиною та застосовувались до дробових октавних комплексно-згладжених спектрів. Потім еквалайзер отримували за допомогою інверсії або алгоритму Кіркебі [81], або аналізу LPC, з мінімальним фазовим вирівнюванням. У розглянутих умовах середнє середнє дало найкращі результати, інші методи також забезпечували подібні показники. Підхід до вилучення прототипу, заснований на середньому середньому, також був поєднаний із методом [73] та застосований до вирівнювання реакції приміщення [75,145]. Суб'єктивні тести на аудіювання підтвердили хороші результати, отримані підходом [75]. Цей підхід був додатково розширений у [146,147], також розглядаючи вирівнювання затримки групи. У роботі [146] після визначення мінімального фазового еквалайзера згладжені фазові відгуки, виміряні в різних положеннях, коригуються з фазовою характеристикою еквалайзера і використовуються для

визначення групових реакцій затримки. Групова затримка прототипу обчислюється шляхом усереднення групової затримки в різних положеннях, а після спектрального згладжування використовується для вилучення еквалайзера групової затримки FIR для всього проходу. У роботі [147] прототип фазової реакції, що використовується для визначення фазового еквалайзера, витягується лише з ранніх відбиттів, які представляють внесок прямого звуку, відкидаючи пізні відбиття, що представляють відбиття середовища. Час змішування між ранніми та пізніми відбиттями обчислюється з використанням підходу, представленого в [147,148] на основі оцінок Гауса. Функція прототипу скорочується з використанням часу змішування, і отримується фазовий еквалайзер FIR за допомогою техніки відповідного фільтра; тобто, зворотний за часом імпульсний відгук на всі проходи. При такому підході артефактів попереднього дзвінка вдається уникнути, оскільки в еквалайзері враховуються лише ранні відображення. Насправді, беручи до уваги лише перші відображення, враховуються лише основні характеристики приміщення, і уникаються ті частини імпульсної реакції, які містять нулі, що змінюються залежно від положення, і згідно з [149] виробляють артефакти попереднього дзвінка.

Прототипний підхід також застосовується в [93,150–153]. На думку авторів, „частина впливу кімнати для прослуховування є природною для людського вуха і не повинна усуватись системою корекції приміщення” [93]. Зокрема, відтворення звуку в кімнаті зазвичай спричиняє підвищений рівень звукового тиску на нижчих частотах через нижче поглинання, яке зазвичай виявляється на цих частотах. Оскільки цей ефект є природним для людського вуха, оскільки він забезпечує відчуття перебування в кімнаті, не слід дозволяти системам еквалайзера кімнати усунути плавне збільшення рівня на низьких частотах, що також називається «посиленням кімнати». Приріст приміщення описує, як ефективність приміщення зростає на низьких частотах порівняно з високими [152]. Більше того, відповідь прототипу повинна зберігати основні характеристики гучномовця; тобто еквалайзер не повинен намагатися "зробити так, щоб усі гучномовці звучали однаково". Таким чином, розроблена система

оцінює основні характеристики гучномовця: нижня гранична частота та нахил, чутливість, індекс спрямованості та верхня гранична межа для динаміка високих частот. Еквалайзер розроблений шляхом отримання інформації як про місцеві властивості в положенні прослуховування, так і про акустичну потужність у тривимірному звуковому полі. RRE заснований на вимірюванні звукового тиску в положенні прослуховування та принаймні в трьох інших випадково вибраних положеннях. Вимірювання в положенні прослуховування містить інформацію про сприймане звукове поле, тоді як інші вимірювання в кімнаті містять інформацію про енергію в тривимірному звуковому полі. Потім інформація використовується для обчислення нижнього та верхнього меж коефіцієнта посилення для розробленого еквалайзера. Відповідь прототипу автоматично обчислюється на основі вимірювань. На низьких частотах відгук прототипу розроблений таким чином, щоб забезпечити той самий коефіцієнт посилення приміщення слухальної кімнати, що відповідає стандарту IEC 268–13 [154], наближаючи плавний коефіцієнт посилення приміщення за допомогою функції стелажу другого порядку, яка плавно додає рівень 6 дБ нижче 120 Гц [152]. Еквалайзер є мінімальною фазою і розроблений на основі гомоморфної техніки [151].

6.4. Загальна компенсація акустичних полів

На низьких частотах у відповіді кімнати можуть з'являтися сильні резонанси. Ці резонанси часто не залежать від положення і пов'язані з тривалими повільно занепадаючими режимами. Для компенсації низькочастотної характеристики були запропоновані різні методи. Багато з цих методів використовують багатоточкові вимірювання для визначення спектральних властивостей резонансів.

Модель RTF із використанням загальних акустичних полюсів, що відповідають резонансам кімнати, запропонована в [63]. Загальні акустичні полюси оцінюються як загальні значення полюсів багатьох низькочастотних RTF, оцінених для різних положень джерела та приймача. Полюси обчислюються з моделі LPC відгуку приміщення, оціненої двома можливими

методами: (i) з використанням методу найменших квадратів, припускаючи, що всі виміряні RTF мають однакові коефіцієнти LPC та (ii) усереднення коефіцієнтів LPC, оцінених з кожного виміряного RTF. Розрахункові полюси відповідають основним резонансним частотам кімнати. Потім, використовуючи оцінені загальні полюси, метод [63] моделює RTF з різними коефіцієнтами ковзного середнього. Модель називається авторами загальною акустичною полюсом та нульовою моделлю, оскільки це модель нульового полюса, утворена загальними акустичними полюсами та нулями, що забезпечуються коефіцієнтами ковзного середнього. Пізніше цей підхід був розширений у [64,65]. У [64] запропоновано багатоточковий фільтр вирівнювання, що використовує загальні акустичні полюси. Спільно-акустичні полюси знову оцінюються як загальні коефіцієнти LPC при багаторазовому вимірюванні RTF. Потім вирівнювання досягається за допомогою FIR-фільтра, що має зворотні характеристики загальної акустичної полюсної функції. Як і для інших повнополюсних моделей, фільтр вирівнювання є мінімально-фазним еквалайзером, який не може компенсувати виїмки частотної характеристики зміни положення приймача. У [65] до загального акустичного еквалайзера полюсів додано етап попереднього кондиціонування. Етап попередньої підготовки пригнічує низькоякісні резонанси у внутрішньому спектрі, тоді як другий етап, заснований на загальних акустичних полюсах, пригнічує або мінімізує низькочастотні резонанси. У [155] запропоновано емпіричний прийом вибору відповідного порядку для загальної моделі акустичного полюса. Методика вибирає перший порядок, для якого подальший ріст не призводить до поліпшення точності моделювання принаймні для одного з вимірюваних RIR. Порядок моделі залежить від обраної максимальної частоти змодельованих полюсів. Ітераційний алгоритм роботи [156] також базується на загально-акустичній та нульовій моделях. Він розробляє біквадратичні фільтри, придатні для багатоточкових RRE.

Загальна компенсація акустичних полюсів також може отримати вигоду від техніки набору фільтрів, заснованої на розділах другого порядку [112–115],

використовуючи, зокрема, роздільну здатність логарифмічної частоти та можливість налаштування полюсних положень.

6.5. Тональна еквалізація

Також пропонується тональна корекція на низьких частотах [4,157]. Тональне вирівнювання спрямоване на контроль надмірно тривалих затухань у слухальних кімнатах, спричинених низькочастотними режимами, при мінімізації чутності цих резонансів. Тональне вирівнювання врівноважує швидкість затухання низькочастотних режимів, щоб відповідати часу реверберації на середній та високій частотах. Це не є технікою RRE як такою, але її можна використовувати із звичайною корекцією амплітуди для оптимізації якості відтвореного звуку. У роботі [157] запропоновано два методи реалізації активної тональної корекції. Перший підхід розглядає один гучномовець і фільтрує звук таким чином, щоб контролювати швидкість затухання режимів (наприклад, використовуючи фільтр з нульовими парами, вирівняними за полюсами, відповідальними за резонанси). Другий підхід реалізує модальне вирівнювання з одним або кількома вторинними гучномовцями. Фільтр еквалайзера перевіряється для кожного додаткового динаміка для отримання компенсаційного звуку. Перший підхід детально вивчений, представлені різні методи визначення режимів, оцінка їх параметрів та проектування еквалайзера. Оцінка параметрів модального демпфування базується на нелінійній оптимізації моделі експоненціального демпфування плюс стаціонарний рівень шуму, представлений у [158].

6.6. Підхід плоскої хвилі

Ще однією можливістю вирівнювання звуку в області низьких частот є та, яку пропонує підхід плоскої хвилі. У прямокутних кімнатах із симетричним розташуванням гучномовців у двох протилежних стінах теоретично показано [159] та експерименти [160], що вирівнювання всередині внутрішньої кімнати можна досягти на низьких частотах. Підхід генерує плоску хвилю, яка поширюється від однієї стіни до протилежної, де вона поглинається гучномовцями. В експериментах [160] сигнали, що подаються на гучномовці,

визначаються з використанням підходу RRE [89]. Датчики помилок розміщені в двох площинах, перпендикулярних до напрямку поширення імітованої плоскої хвилі. Бажаним сигналом у площинах є дельта-функція Дірака із затримкою, що відповідає часу, який потрібно звуку, щоб пройти відстань між площинами. Підхід плоскої хвилі також вивчався авторами [161–165]. По-перше, в [161] автори розробили додаток на основі наближення кінцево-різницевої часової області для вивчення низьких частот при відтворенні звуку. Зокрема, прямокутна кімната змодельована за допомогою дискретної моделі в часі та просторі. Потім у [162] додаток був використаний для вивчення різних конфігурацій гучномовців у кімнаті, щоб зменшити ефект від акустичних режимів. Показано, що за рахунок збільшення кількості гучномовців покращується варіація реакції приміщення між положеннями за рахунок збільшення приросту відхилення величини в кожному положенні. Додаток також використовувався для оцінки ефекту різних технік вирівнювання, таких як багатоточкове вирівнювання та вирівнювання потужності акустичного випромінювання гучномовця. Врешті-решт, рішення для вирівнювання низькочастотного звукового поля за допомогою кількох гучномовців - з назвою керована акустична басова система (CABS) - було запропоновано та вивчено в [163–165]. Це рішення створює рухому плоску хвилю в одній стороні кімнати і відмінняє розташування на протилежній стіні за допомогою додаткових гучномовців із уповільненою та антифазною реакцією для усунення відблисків задньої стінки. Використовуючи застосування [161] та реальні вимірювання в прямокутних кімнатах, автори показали, що розчин CABS може створювати рівномірне акустичне поле в низькочастотному діапазоні. У [166] підхід [159, 163] додатково поширюється на приміщення довільної форми з кількома гучномовцями, «розташованими в більш нормальних місцях», враховуючи налаштування гучномовця 5.0. Крім того, [167] розглядав проблему не прямокутної кімнати та асиметричного налаштування гучномовця. У [167] техніка вирівнювання з кількома входами / кількома виходами (MIMO), яка прописує лише величину реакції приміщення в запропоновано контрольні точки.

Цей підхід дозволяє отримати відхилення меншої величини порівняно з попередніми підходами з плоскою хвилею. Для поліпшення вирівнювання площинними хвилями в [168] було запропоновано підхід управління, який називається регуляцією варіації зусиль. У цьому підході звичайна функція витрат RRE_{of} [169] - заснована на мінімізації похибки найменших квадратів у декількох контрольних точках - модифікується додаванням терміну регуляризації, пропорційного квадратичним відхиленням між силою джерела. Підхід може застосовуватися як у частотній та часовій областях. Результати моделювання показують, що методика може призвести до менших глобальних помилок відтворення та кращих показників вирівнювання при розташуванні позицій подалі від контрольних точок, ніж регуляризація Тихонова або підхід, заснований на подачі одного і того ж сигналу на всі гучномовці, розміщені на одній стіні.

6.7. Інші низькочастотні підходи RRE

На дуже низьких частотах замість плоскої хвилі набагато ефективніше використовувати підхід під тиском і полем камери [170]. Цей підхід отримується шляхом надсилання одного і того ж сигналу на всі гучномовці, що генерує всередині кімнати візерунок стоячої хвилі, який є однорідним на довжинах хвиль, значно більших за кімнату. З цієї причини в [170] пропонується підхід відтворення гібридного поля, який поєднує ефективність відтворення поля тиску на дуже низьких частотах з однорідним звуковим полем, отриманим при наближенні плоских хвиль на більш високих частотах. У [79] проблема багаточастотного гучномовця низькочастотного RRE для широкої зони прослуховування, причому еквалізований гучномовець підтримується рештою, розглядається як проблема багатомікронного мінімізації між бажаною реакцією та синтезованою величиною відгуку. фізичні та психоакустичні критерії. Зокрема, для отримання коротких фільтрів вирівнювання на фільтри вирівнювання накладається обмеження тимчасового маскування. Щоб уникнути відчутних відлунь, розглядається комбінація затримки та посилення щодо головного гучномовця, при цьому сигнали допоміжних гучномовців повинні

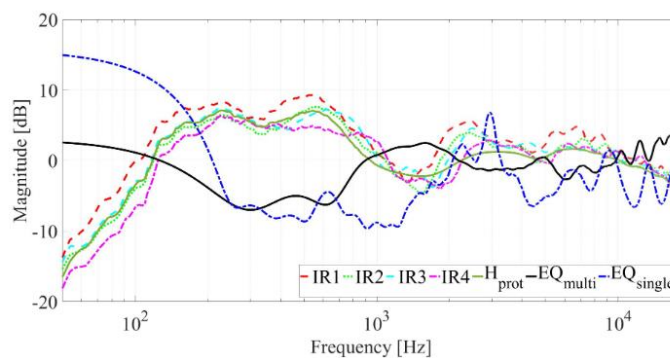
падати нижче порога відлуння [171]. Щоб уникнути модифікацій просторового сприйняття, затримка додаткових сигналів гучномовця змушена бути щонайменше 1 мс, щоб використати ефект переваги. Щоб уникнути збільшення надрізів, максимальний коефіцієнт підсилення накладається на еквалайзери. Фільтри вирівнювання приміщення обчислюються з урахуванням опуклої системи оптимізації, яка враховує всі ці обмеження.

6.8. Квазіанехогенний підхід

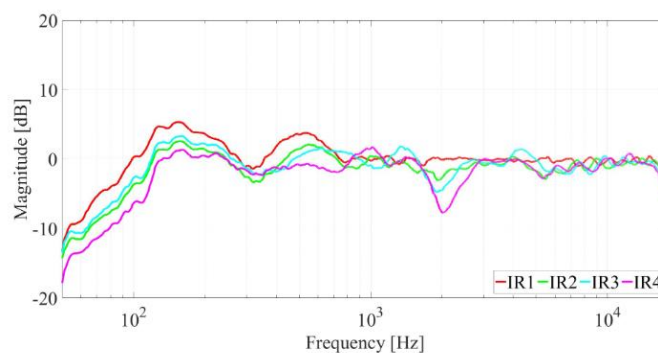
Підходом, який доповнює низькочастотні методи, представлені в попередньому підрозділі, є квазіанехогенний підхід [172]. На середніх і високих частотах у тембровому сприйнятті та локалізації переважає прямий звук. Таким чином, в [172] квазіанехоїчна реакція гучномовця отримується як закрита версія (до першого відображення) RIR і використовується для проектування еквалайзера у два етапи. По-перше, змішаний фазовий еквалайзер отримують із квазіанехогенного RIR, обчислюючи зворотний фільтр із підходом найменших квадратів. Квазіанехогенна реакція гучномовця має невелику тривалість, а затримка, введена еквалайзером, занадто коротка, щоб отримати попередні сигнали дзвінка. Потім використовується мінімальний фазовий еквалайзер для корекції залишкової частини відгуку приміщення (тобто, модифікації спектра величини, викликані реверберацією).

У роботі [173] квазіанехогенний підхід поєднується з прототипним підходом, описаним у Розділі 6.3. Зокрема, нова функція прототипу походить від поєднання квазіанехогенних імпульсних реакцій з імпульсними відгуками, записаними в реальному середовищі, для вирівнювання. Підхід використовується для вирівнювання прямого звуку лише в середньочастотному діапазоні, застосовуючи при цьому повне вирівнювання в модальному діапазоні частот. Підхід знову мотивований тим, що на середніх і високих частотах у сприйнятті та локалізації тембру домінує прямий звук. Таким чином, вимірювані, але переважно нечутні відхилення величини внаслідок відбиттів не слід вирівнювати [174]. У роботі [173] було проведено кілька експериментів з метою підтвердження запропонованого підходу, звітування про об'єктивні

вимірювання та суб'єктивні тести прослуховування в порівнянні з підходами сучасного рівня. У цьому контексті на рисунках 8 представлені результати процедури вирівнювання. Зокрема, на малюнку 8а показані чотири імпульсні характеристики, отримані в ареальній кімнаті - функція прототипу та еквайзер, отримані за допомогою багатоточкового підходу [173], і одноточковий еквайзер, отриманий як зворотний фільтр згладженої частотної характеристики IR1. На малюнку 8b показано вплив процедури вирівнювання на ІЧ, що застосовує багатоточковий підхід, тоді як на рис. 8в - ефект одноточкового еквайзера. Очевидно, що ефективність одноточкового еквайзера дуже хороша лише для IR1, тоді як багатоточковий еквайзер демонструє більш рідкісні реакції порівняно з результатами, отриманими при одноточковому підході.



(a)



(b)

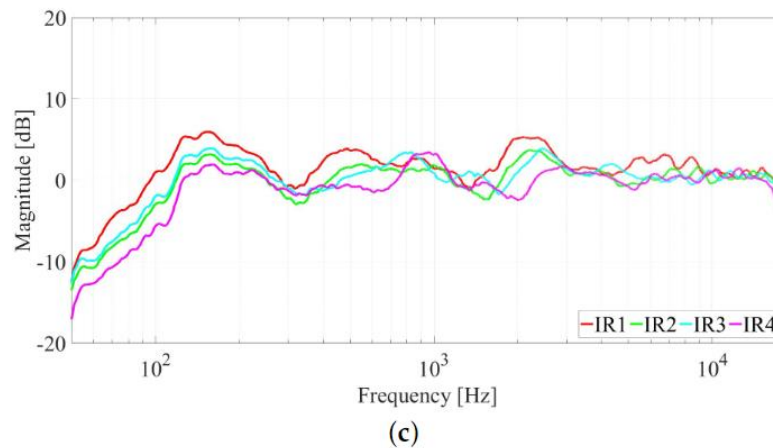


Рис 8. Процедура вирівнювання: (а) імпульсні відгуки (ГЧ) з функцією прототипу H_{prot} , багатоточковий еквалайзер EQmulti, одноточковий еквалайзер EQsingle. Результати вирівнювання із застосуванням (b) одноточкового еквалайзера EQsingle; та (c) багатоточкового еквалайзера EQmulti.

7. Адаптивне одноточкове та багатоточкове вирівнювання

Кімната, як правило, змінюється в часі («слабо нестационарна» система, як визначено в [2]), яка змінюється в залежності від кількох параметрів, таких як положення фізичних предметів у кімнаті, відкривання дверей, а також як пересування людей та інші перешкоди в огороженні [6,175]. Крім того, коливання температури можуть призвести до значних коливань RIR, як повідомляє [176]. Крім того, можуть спостерігатися зміни в положеннях джерела та приймача, а також характеристиках гучномовця та мікрофона, як повідомляється в [6]. Таким чином, слід прийняти адаптивні рішення, придатні для відстеження та виправлення повільних змін у реакції приміщення. У літературі пропонуються різні адаптивні методи RRE. Тут підходи класифікуються з урахуванням кількості вхідних та вихідних каналів як SISO / SIMO та MISO / MIMO, де вхід відноситься до кількості гучномовців, а вихід до кількості мікрофонів, оскільки ці класи мають подібні проблеми в процедурі ідентифікації.

7.1. Підходи SISO / SIMO

Ці методи можна класифікувати за підходами до часової та частотної областей.

7.1.1. Підходи до часового домену

Перший адаптивний еквалайзер був запропонований у [77], враховуючи мінливість середовища з різних поглядів. Підхід базувався на одноточковій техніці, адаптивно мінімізуючи у часовій області середньоквадратичну помилку між вирівняною реакцією та відстроченою одноканальною версією вихідного сигналу з використанням алгоритму х-фільтрації. Вирівнювання було ефективним для розглянутої позиції, але було погіршення стану в інших точках огороження, що також описано в Розділі 6. Тому багатоточковий підхід був також представлений тими ж авторами в [77], де еквалайзер був розроблений шляхом адаптивної мінімізації суми квадратичних помилок між вирівняними відповідями в декількох положеннях і відкладеною версією вхідного сигналу. На жаль, підхід дуже чутливий до піків та виїмок у реакції приміщення та до варіацій відгуку кімнати в різних положеннях. Як наслідок, проблеми з ранніми відлуннями можуть бути легко відчутні.

7.1.2. Підходи до частотних доменів

При роботі в частотній області одноточковий метод вимірювання відгуку приміщення був запропонований в [177]. Тут сигнали гучномовця та мікрофона поділяються на піддіапазони (20-смуговий фільтр-банк), а вирівнювання досягають шляхом адаптивного оновлення ваги відповідного фільтру в цих піддіапазонах. Підхід цікавий тим, що поєднує в собі простоту, стійкість до піків та виїмок реакції приміщення, а також можливість відстежувати варіації реакції кімнати. Це було вдосконалено у [178] введенням частотно-залежного розміру кроку. Таким чином можна оптимізувати адаптивне вирівнювання в кожному піддіапазоні, покращуючи загальну швидкість конвергенції. У [179] також було представлено подальше вдосконалення попередніх методів [177,178], щоб упоратися з онлайн-ідентифікацією імпульсної реакції. Зокрема оцінку відгуку приміщення отримують вставляючи з маскуванням частот штучних тестових сигналів таким чином, що слухачі їх не чують. Потім сигнал аналізується в частотній області для ідентифікації тестового сигналу та визначення імпульсної характеристики приміщення. У [180] підхід [177,178] був розроблений та вдосконалений шляхом розробки багатоточкового рішення. Після ідентифікації

в смугах частот до імпульсних характеристик застосовують дробове згладжування октави, а фільтр-прототип обчислюється із середнього значення величини відгуку кімнати. Отримані результати показали, що ефективність цієї досить простої конструкції можна поліпшити, розглянувши багатоточкове рішення, що призводить до збільшення ширини зрівняльної зони. У [84] підхід [180] був додатково розроблений з урахуванням перекосу частоти в області низьких частот для поліпшення сприйняття. Зокрема відгуки кімнати в різних положеннях вирівнюваної зони оцінюються в деформованому домені, і загальна тенденція цих відповідей витягується як функція-прототип. Це дозволяє збільшити роздільну здатність еквайзера на частотах, де слухова система людини є більш чутливою. Адаптивні версії методів набору фільтрів [112–115] також можуть бути використані з тією ж метою.

7.2. Підходи MISO / MIMO

Адаптивні методи вимірювання відгуку приміщення, запропоновані в [84,178,180] (та багатьох інших роботах), розглядають питання еквалізації одного джерела звуку (тобто одного каналу відтворення звуку) через проблему оцінки декількох імпульсних реакцій одночасно. При використанні двох або більше каналів коваріаційна матриця багатоканального адаптивного алгоритму стає погано обумовленою через кореляції між каналами для типових методів відтворення. Порушення кондиції, як правило, викликає проблеми зближення. Це було показано, наприклад, для стереофонічного акустичного приглушення ехо [181]. Щоб впоратися з проблемою неунікальності, зазвичай використовують метод зменшення когерентності між каналами. У цьому контексті багато методів, які використовують для зменшення перехресних кореляцій каналів, часто вносять значні спотворення, і це неприйнятне у високоякісних системах відтворення звуку [181,182]. Отже, слід розглянути підгожу техніку, яка може декорелювати сигнали гучномовця та зберегти якість звуку. Підхід у [183] вводить багатоканальне рішення, яке також розглядає проблему неунікальності. Відгуки в приміщенні оцінюють з хорошою точністю, зменшуючи когерентність між каналами за допомогою техніки, що забезпечує

лише незначне погіршення якості звуку. Зокрема низькочастотна область декорельована шляхом використання відсутнього фундаментального явища, тоді як високі частоти декорельовані за допомогою всепропускного фільтру другого порядку в поєднанні з мультирежекторним [184]. Еквалайзер призначений міняти частотну область, щоб поліпшити спектр в області низьких частот і одночасно зменшити обчислювальні витрати на конструкцію. У роботі [185] адаптивна багатоканальна і багатопозиційна система вимірювання відгуку приміщення, що коротко представлена в [183], детально розширена і гарантує вже сьогодні реалізацію комерційних продуктів в галузі музичної техніки.

Щоб поліпшити швидкість збіжності та надійності алгоритму адаптивної ідентифікації за наявності низького відношення сигнал/шум нещодавно в [186] було запропоновано користувати розпізнавальний адаптивний алгоритм для системи MIMO (multiple-input/multiple-output — багатоканальний вхід / багатоканальний вихід). Алгоритм базується на вдосконаленому пропорційному нормованому алгоритмі найменших середніх квадратів (англ. IPNLMS) в рамках звичайної схеми відфільтрованого x (англ. IPNLMS-FX), раніше запровадженої для активного контролю шуму (АКШ) [187], і тут розширений багатоканальне вирівнювання. Однак цей метод вимагає апіорної оцінки імпульсних характеристик, яка недоступна у багатьох практичних застосуваннях. З тією ж метою покращення збіжності та стійкості у [188] була запропонована комбінація адаптивних фільтрів на основі блоків (також із використанням розпізнавальних алгоритмів).

Варто підкреслити: якщо розглядати бінауральну систему, маємо природню декореляцію між стереоканалами. Стереоподання адаптивної системи вимірювання відгуку приміщення може бути досягнуте без декореляції каналу, що стверджують у [169,189]. Удосконалення цієї техніки подано в [190], де пропонують структуру піддіапазону для зменшення обчислювальної складності процедури.

8. Фіксоване та адаптивне вирівнювання хвильового домену

Досі розглянуті підходи еквалізації враховували відтворене звукове поле в одній або декількох точках простору. Ці точки в ідеалі мають збігатися з потенційним положенням слухача або обмеженою зоною прослуховування. Ширший погляд на еквалізацію можна отримати, якщо врахувати все відтворене звукове поле в межах бажаної (потенційно великої) зони прослуховування. Цього можна досягти, враховуючи просторово-часовий характер звукового поля замість звукового тиску в обмеженій кількості точок. Перехід від точок до поля ґрунтуватиметься на низці засад у цьому підрозділі. Це буде подано як огляд відповідних джерел.

8.1. Фізичний фон

Інтегральне рівняння Гельмгольца (ІРГ) [191] забезпечує розв'язок неоднорідного хвильового рівняння для однорідних граничних умов. Це охоплює зокрема й звукове поле, відтворене розподілом динаміків у кімнаті. Згідно з ІРГ звуковий тиск у вільному об'ємі джерела та розсіювача V однозначно визначається звуковим тиском та його спрямованим градієнтом на межі гучності ∂V . Це може бути використано для аналізу звукових полів, а також для їх синтезу. Для цього достатньо зафіксувати звуковий тиск та його градієнт на межі потрібного об'єму. Те саме стосується синтезу звукових полів, коли розміщення гучномовців навколо зони прослуховування дозволяє повністю контролювати звукове поле в цій зоні. Однак з погляду технічної складності, як правило, не бажано фіксувати що звуковий тиск, що його градієнт спрямованості за допомогою двох різних розміщених на межі зони прослуховування типів мікрофонів. Це стосується і синтезу за допомогою гучномовців. Тут потрібно було б використовувати монопольні та дипольні гучномовці. Мікрофони й динаміки з властивостями монополя бажаніші за їхні дипольні аналоги. У [192] показано, що ІРГ можна звести до монопольного варіанту за певних практично можливих умов. Це закладає теоретичну основу вимірювання відгуку приміщення в розширеній зоні прослуховування. Таким чином звукове поле в зоні прослуховування можна аналізувати та контролювати за допомогою безперервного розподілу розташованих на межі цієї зони

Для практичної реалізації вищезазначених принципів можна використовувати лише обмежену кількість мікрофонів та гучномовців. Отже, безперервний розподіл мікрофонів та гучномовців повинен бути відібраний просторово. Геометрія та вибірка проілюстровані на рисунку 9.

The diagram illustrates a listening room with a central shaded region labeled "listening area". The room boundary is marked by a series of discrete points, with the first point labeled 0. The boundary is divided into two segments: an inner segment labeled 1 and an outer segment labeled 2. The total number of points on the boundary is labeled N . A point on the inner boundary is labeled M . A point on the outer boundary is labeled n . A point on the inner boundary is labeled m . A point on the outer boundary is labeled V . The label "discrete synthesis and analysis positions" points to the boundary points. The label "listening room" is at the bottom left.

8.2. Доменно-хвильове адаптивне фільтрування

210

змінюватися (в кімнату зайдуть люди, зміниться температура абощо). Наприклад, наслідки зміни кімнатної температури при використанні статичних фільтрів проілюстровані в [198]. У минулому було розроблено найрізноманітніші алгоритми адаптації. Оскільки вимірювання відгуку приміщення є оберненою задачею, клас алгоритмів х-фільтрації добре підходить. Фільтри можуть бути обчислені адаптивно за допомогою багатоканального алгоритму х-фільтрування рекурсивного найменших квадратів (X-RLS) [199]. Однак у контексті багатоканального вимірювання відгуку приміщення адаптивне рішення має три основні проблеми: 1) погана кондиціонованість; 2) неунікальність; 3) числова складність. Перша проблема пов'язана з просторово-часовою кореляцією типових сигналів гучномовця, друга з основною проблемою оптимізації, а третя з розміром типових систем МІМО, що відповідає теоретично-хвильовому погляду. Розв'язувати третю проблему, яка також доповнює інші дві, пропонують за допомогою адаптивного фільтрування доменів хвиль (АФДХ) [10 200]. Тут основна система МІМО розділена набором просторово-часових перетворень, як показано на рис. 10.

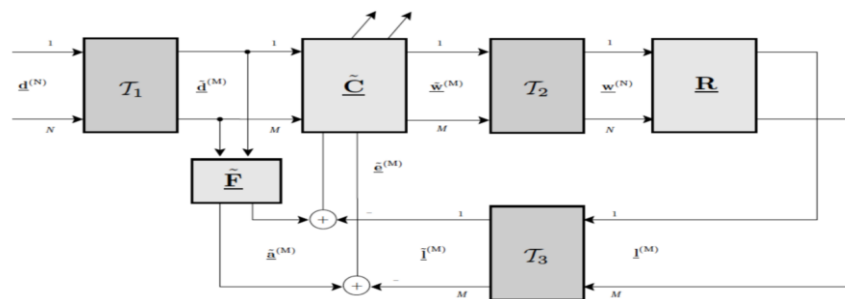


Рис. 10. Блок-схема, що ілюструє концепцію еквалізації приміщення на основі доменно-хвильової фільтрації (ДХФ). Управні сигнали для N гучномовців (позначені $d^{(N)}$) перетворюються на M хвильових доменів $\tilde{d}^{(M)}$ за допомогою просторово-часового перетворення T_1 . Вони фільтруються в хвильовій області за допомогою матриці МІМО \tilde{C} фільтрів вирівнювання, в результаті чого отримують попередньо відфільтровані сигнали керування гучномовцем $\tilde{w}^{(M)}$, які потім трансформуються назад за допомогою T_2 . Акустичні шляхи між N гучномовцями та M контрольними точками (мікрофонами) об'єднані в МІМО-матрицю передачі кімнати R . Сигнали в контрольних точках $l^{(M)}$ перетворюються на хвильові домени управних сигналів $\tilde{l}^{(M)}$ за допомогою перетворення T_3 . Потрібне розповсюдження у відкритому полі моделюється в хвильовій області за допомогою МІМО-матриці \tilde{F} передавальних функцій відкритого поля як шукані сигнали $\tilde{a}^{(M)}$ в контрольних

точках. Похибка $\tilde{\epsilon}(M)$, яка використовується для адаптації компенсаційних фільтрів, визначається різницею перетворених шуканих сигналів $\tilde{a}(M)$ та фактичних сигналів $\tilde{f}(M)$ у контрольних точках.

Перетворення $T1$ в $T3$ мотивовані фізичними передумовами задачі еквалізації кімнати та її розв'язання з використанням ортогональних розширень, як зазначено в попередньому розділі. З точки зору основної багатоканальної проблеми, це може бути досягнуто діагоналізацією систем МІМО з використанням узагальненого розкладання особливих значень (УРОЗ). Цей підхід відомий як адаптивна фільтрація власного простору (АФВП) [193]. Як наслідок, проблема адаптації зводиться до адаптації основних діагональних елементів фільтра еквалізації кімнати МІМОС у трансформованій області. Таким чином, обчислювальна складність значно знижується, а проблема неунікальності зменшується. Однак УРОЗ вимагає, щоб були відомі шляхи передачі від гучномовців до мікрофонів, що суперечить ідеї адаптивного обчислення фільтрів еквалізації. Використовуючи засновані на вільних полях хвильового рівняння аналітичні перетворення, було досягнуто приблизну діагоналізацію системи МІМО [10,200].

Оригінальний підхід був зосереджений на адаптації лише діагональних шляхів у трансформованій області. У [201] це було розширено до гнучкої системи адаптації, що також враховує недіагональні шляхи. Повна адаптація всіх шляхів у трансформованій області досліджена в [202]. Зворотні перетворення для АФДХ були введені в [203], тоді як піддіапазонний підхід до АФДХ був опублікований в [204]. Крім того стратегії використання масивів гучномовців з нерегулярним інтервалом були запропоновані в [205].

8.3. Трансформовані підходи домену

АФДХ використовує набір перетворень, які перетворюють багатоканальну проблему адаптивного вирівнювання в трансформований домен. Ця основна ідея застосування просторової трансформації також була застосована до неадаптивного вирівнювання приміщення, спрямованого на велику зону прослуховування. У [206] звукове поле був розкладений на кругові / циліндричні

основні функції для концентричного налаштування гучномовця та мікрофонної решітки. Це, по суті, двовимірна проблема. Фільтри вирівнювання були розраховані шляхом оптимізації найменших квадратів у перетвореному домені. Вирівнювання кімнат також розглядалося в контексті багатозонного синтезу шляхом формулювання тривимірної задачі сферичної гармоніки (наприклад, [207]). Досить інший підхід розглядається в [208]. Тут оригінальний НІЕ інтерпретується таким чином, що зовнішнє звукове поле сферичного гучномовця послаблюється використанням гучномовців зі змінною спрямованістю. Послаблення зовнішнього звукового поля призводить до зменшення відблисків, що повертаються в зону прослуховування. Хоча такі гучномовці ще не реалізовані, результати моделювання виглядають перспективними. Проблема вирівнювання розглядається у сфері сферичних гармонік, де фільтри обчислюються шляхом оптимізації найменших квадратів.

8.4. Методи з урахуванням геометрії кімнати

Знання геометрії кімнати можна використовувати для обчислення результуючого звукового поля в кімнаті, наприклад методом дзеркального зображення. Потім можливості керування системою синтезу звукового поля можуть бути використані для скасування передбачуваних внесків із кімнати. Методи, які явно використовують знання з геометрії кімнати, можна розглядати як спеціалізацію обговорюваних до цього часу методів, оскільки вони базуються на теоретично-хвильовому погляді на проблему. Метод вирівнювання ранніх відбиттів для синтезу хвильового поля (WFS) опублікований в [209]. Тут джерела дзеркальних зображень анулюються антифазними джерелами віртуальних точок, розміщеними в заздалегідь обчислених положеннях джерел зображення. Подібний підхід представлений у [210] для амбісонік вищого порядку. Підхід до вирівнювання приміщення для лінійного гучномовця, що створює пучки для віртуальної системи об'ємного звучання, обговорюється в [211]. Вирівнювання кімнатних відбиттів досягається обліком відображення променів у приміщенні. Фільтри вирівнювання обчислюються шляхом вирішення основної задачі найменших квадратів у закритій формі. У роботі

[212] представлений метод, який базується на чисельному моделюванні імпульсних реакцій між гучномовцями та контрольними точками. Розглядаються лише ранні роздуми. Імітовані імпульсні відгуки подаються в МІМО-вирішувач для виведення фільтрів вирівнювання.

8.5. Підходи МІМО та SIMO

Як альтернатива теоретично-хвильовому підходу, що обговорювався до цього часу, акустичні шляхи між гучномовцями та мікрофонами можна інтерпретувати як незалежні лінійні інваріантні в часі системи. Всі отримані функції передачі можуть бути об'єднані разом у систему з декількома входами / з декількома виходами (МІМО). Підходи МІМО для вирівнювання приміщень відрізняються, серед іншого, стосовно положень гучномовця та мікрофона (контрольних точок), а також конкретної техніки, що використовується для обчислення фільтрів вирівнювання. Різниця між теоретико-хвильовим та МІМО-підходами, обговореними в подальшому, полягає в тому, що обчислення фільтрів вирівнювання не виконуються в просторово перетвореній області. Хоча розміщення гучномовців та мікрофонів на межі зони прослуховування мотивується ІРГ, підходи МІМО можуть відхилятися від цього розташування. Як зазначено вище, для синтезу та захоплення всього звукового поля до заданої частоти необхідно використовувати достатню кількість динаміків та мікрофонів. Якщо дискретизація недостатньо щільна, вирівнювання може бути досягнуте лише в місцях мікрофона або поблизу них.

Неадаптивний підхід МІМО, який безпосередньо виникає внаслідок дискретизації ІРГ, представлений у [213]. Система МІМО інвертована для обчислення фільтрів вирівнювання для глобального вирівнювання. В якості альтернативи також обговорюється місцеве рішення. Подібний підхід дотримується в [85] щодо синтезу хвильового поля. Скорочення каналів також досліджувалося в контексті вирівнювання МІМО [214,215] на основі рішення найменших квадратів.

Обчислення фільтрів вирівнювання, як правило, є оберненою проблемою. Запропоновано різні алгоритми, які покращують чисельну та обчислювальну

ефективність, а також чисельну обумовленість - наприклад, швидкий ітераційний алгоритм інверсії MIMO, що працює в DFT-домені [216], або наближення DFT-домену MIMO, відфільтрованого - х алгоритм [217]. У роботі [216] представлений алгоритм найвищого спуску та наближений Гаусс – Ньютон для проектування еквалайзера MIMO. У роботі [218] запропоновано метод боротьби з низьким кондиціонуванням матриці передавальної функції на деяких частотах. Проблема виправлена шляхом вивчення структури матриці функції передачі MIMO та заміни її оберненої матриці на псевдообернену, що забезпечує цілий ряд прийнятних рішень. Поліноміальні формулювання MIMO проблеми вирівнювання приміщення обговорюються в [219,220] з розширеннями, спрямованими на явний контроль кількості активних гучномовців, що використовуються для вирівнювання [221].

Існує також ряд спеціалізованих підходів до вирівнювання для конкретних сценаріїв. Наприклад, вирівнювання багатоканальних стереофонічних систем за умови, що стереофонічні пари гучномовців повинні мати подібні передавальні функції, обговорюється в [222–224]. Підхід розділений на два етапи: (I) вирівнювання одного тракту, також використовуючи інші гучномовці, та (II) оптимізація подібності між двома каналами, які використовуються для стереофонічного зображення. Вирівнювання кімнати в автомобілях розглядалося в різних дослідженнях. Неадаптивний підхід вирівнювання MIMO із використанням фільтрів IIR або FIR представлений у [225]. Оптимізація проводиться з точки зору загальної величини амплітуди, щоб уникнути проблем із забарвленням / тонами. Комбінований підхід до вирівнювання приміщення та скасування перехресних розмов для автомобілів обговорюється в [226].

Окрім розглянутих дотепер підходів MIMO, також були досліджені підходи до вирівнювання одиниць гучномовців, які використовують кілька мікрофонів. Це становить проблему з одним входом / кількома виходами (SIMO). Неадаптивний поліноміальний багатовимірний підхід до управління у поєднанні з обмеженою середньоквадратичною конструкцією помилок та нульовою кластеризацією обговорюється в [149]. Статистичний метод висновку,

який враховує статистичні зміни між різними положеннями мікрофона для поліпшення стійкості та збільшення зони прослуховування, представлений в [227].

9. Методи оцінки RRE

Одним із важливих аспектів є оцінка результатів RRE з урахуванням інструментальних заходів або суб'єктивних тестів на аудіювання. Перший спрямований на заходи, які відповідають меті процедури, наприклад, кількісно визначають подібність між цільовою функцією та результатом вирівнювання. Однак важлива роль повинна відводитися перцептивному оцінюванню, оскільки остаточне судження завжди виконує слухач людини в конкретному середовищі. У цьому розділі ми спочатку аналізуємо інструментальні параметри, що використовуються як основний етап аналізу отриманих результатів. Потім подається огляд найпоширеніших процедур тестування на аудіювання.

9.1. Інструментальні вимірювання

У наступному розділі розглядаються найпоширеніші інструментальні вимірювання для оцінки RRE. В усьому розділі $h(n)$ позначає RIR у дискретному часовому діапазоні, тоді як $H(e^{j\omega})$ позначає своє перетворення Фур'є з дискретним часом, причому ω є нормованою кутовою частотою.

9.1.1. Заходи спектрального відхилення

Спектральне відхилення вперше було використано для оцінки процедури RRE в [76], а потім було прийнято в багатьох інших роботах [92, 228]. Спектральне відхилення частотної характеристики $E(e^{j\omega})$ SD може бути виражене як:

$$S_D = \sqrt{\frac{1}{Q_h - Q_1 - 1} \sum_{i=Q_1}^{Q_h} \left(10 \log_{10} \left| E \left(e^{j \frac{2\pi}{N} i} \right) \right| - D \right)^2},$$

де

$$D = \frac{1}{Q_h - Q_1 - 1} \sum_{i=Q_1}^{Q_h} \left(10 \log_{10} \left| E \left(e^{j \frac{2\pi}{N} i} \right) \right| \right),$$

де Q_l та Q_h - найнижчий та найвищий індекси частоти відповідно зрівняльної смуги. Зазвичай експериментальні результати забезпечують початкове спектральне відхилення $S_{D, in}$, в, розраховане з $E(e^{j\omega}) = H(e^{j\omega})$, і остаточне спектральне відхилення $S_{D, fin}$, обчислене після вирівнювання з урахуванням $E(e^{j\omega}) = H(e^{j\omega}) \cdot H_{inv}(e^{j\omega})$, де $H_{inv}(e^{j\omega})$ являє собою розроблений еквалайзер. На рисунку 11 показані криві, що використовуються для розрахунку S_D . Також було розглянуто середнє спектральне відхилення (MSDM), яке представляє середнє значення остаточного спектрального відхилення за всією сукупністю виміряних RIR [74, 146].

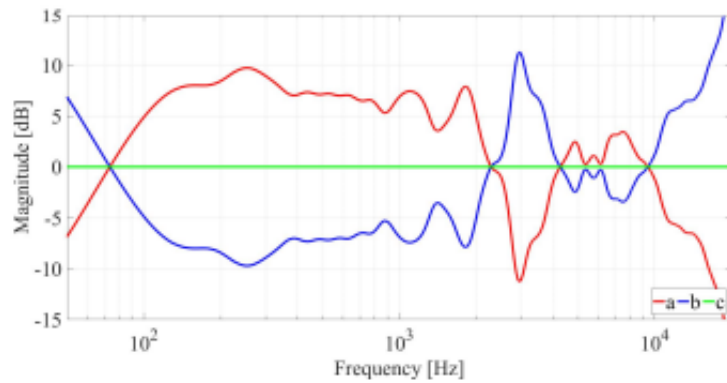


Рис 11. Результати вирівнювання: спектральне відхилення вимірює відстань між вирівняною кривою відгуку і плоскою: (червона) $H(e^{j\omega})$, (синя) $H_{inv}(e^{j\omega})$, (зелена) $E(e^{j\omega}) = H(e^{j\omega}) \cdot H_{inv}(e^{j\omega})$.

Аналогічно середньому показнику спектрального відхилення, який дає міру відхилення частотної характеристики величини від плоскої [228], середній показник відхилення групової затримки був введений в [147, 229], щоб кількісно визначити середнє відхилення в термінах групи затримка:

$$GD_d = \frac{1}{M} \sum_{i=1}^M \sqrt{\frac{1}{Q_h - Q_l - 1} \sum_{i=Q_l}^{Q_h} (GD_1(i) - K_l)^2},$$

де

$$K_l = \frac{1}{Q_h - Q_l - 1} \sum_{i=Q_l}^{Q_h} GD_1(i)$$

Q^l і Q_h - найнижчий та найвищий індекси частоти, відповідно, вирівняної смуги, а $GD_l(i)$ - затримка групи RIR для i -го індексу частоти. Завданням змішаного фазового вирівнювання є досягнення лінійної фази, тому групова затримка повинна бути максимально рівною: використовуючи цей параметр, можна кількісно визначити відстань отриманої групової затримки від постійної затримки.

9.1.2. Карта Саммона

Карта Саммона була введена для оцінки RRE у [230]. Це нелінійний метод проєкції, який відображає багатовимірні дані на менші розміри (наприклад, два або три). Головною властивістю карти Саммона є те, що вона зберігає геометричні відстані між сигналами в багатовимірному просторі у двох-трьох вимірах. Враховуючи M -величину відгуку $|H_k(e^{j\omega})|$, $k = 1, \dots, M$, виміряних RIR, алгоритм карти Саммона ітеративно мінімізує - за схемою градієнтного спуску - сукупну суму різниць між евклідовими відстанями у великому та низькому розмірному просторі. Наступна цільова функція зведена до мінімуму:

$$J_{\text{Sammon}} = \frac{1}{\sum_{l=1}^M \sum_{m=l+1}^M d_{lm}} \sum_{l=1}^M \sum_{m=l+1}^M \frac{(d_{lm} - d'_{lm})^2}{d_{lm}},$$

де

$$d_{lm} = \sum_{w=1}^W \left(\left| H_l \left(e^{j \frac{2\pi}{N} w} \right) \right| - \left| H_m \left(e^{j \frac{2\pi}{N} w} \right) \right| \right)^2,$$

$$d'_{lm} = \sum_{i=1}^L |r_l(i) - r_m(i)|^2$$

де W - кількість однаково розташованих частот, а L - розмірність простору карти Саммона. На карті Саммона точка, пов'язана з $H_k(e^{j\omega})$, представлена як $(r_k(1), \dots, r_k(L))$.

Розглядаючи двовимірне відображення ($L=2$), при збіжності точки $r_k(1)$, $r_k(2)$ з $k = 1, \dots, M$ сконфігуровані на двовимірній площині таким чином, що відносні відстані між різними $H_k(e^{j\omega})$ візуально помітні. Після вирівнювання отриману продуктивність можна визначити за розміром і формою області,

визначеною вирівняними частотними характеристиками на карті. Кругла форма навколо нулів вказує на рівномірне вирівнювання в усіх місцях [230]. На рисунку 12 показані результати, отримані за допомогою карти Саммона: можна помітити, що для ГЧ без вирівнювання (рис. 12а) точки розташовані далеко від центру карти, тоді як для ГЧ з вирівнюванням (рис. 12б) рівномірно розподілені по центру карти.

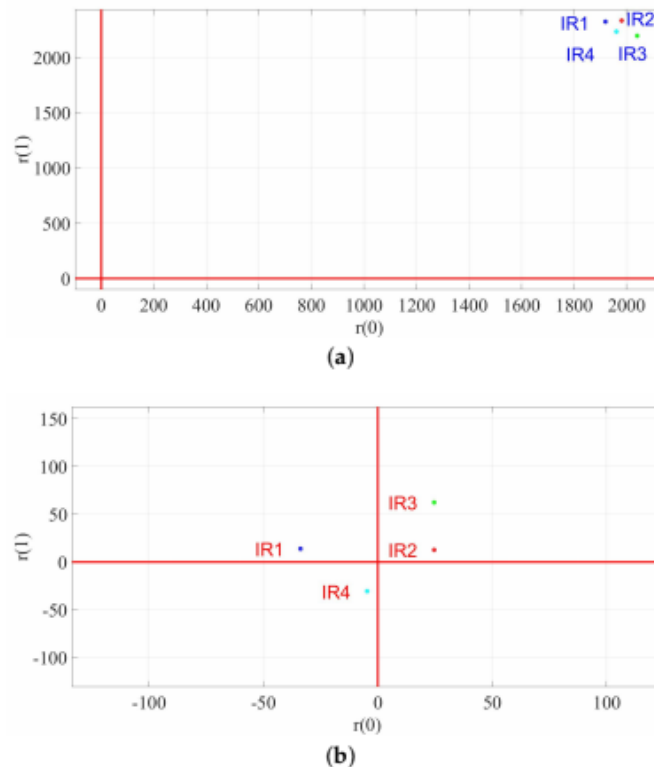


Рис 12. Результати карти Саммона: (а) RIR без вирівнювання; (б) RIR з вирівнюванням

9.1.3. Полегшення енергетичного спаду

Ефект від вирівнювання може бути розрахований (оцінений) виходячи з параметру *energy decay relief* (EDR) (поверхні затухання енергії), що узагальненням кривої(залежності) затухання енергії в координатах часу та частоти, що використовується для розрахунку часу реверберацій $T60$. Оскільки режими кімнати характеризуються піковими значеннями АЧХ та розширеним відгуком в часовій області, величина EDR допомагає зрозуміти ефект отриманий від процедури вирівнювання. EDR визначається як затухання енергії імпульсної характеристики кімнати в координатах час-частота, і на неперервному відрізку часу визначається наступною залежністю:

$$EDR_h(t, f) = \int_t^{+\infty} \rho_h(\tau, f) d\tau$$

де $\rho_h(\tau, f)d\tau$ - енергетичне представлення імпульсної характеристики кімнати в координатах час-частота, отримане застосуванням короткочасного перетворення Фур'є до прямокутного вікна аналізу.

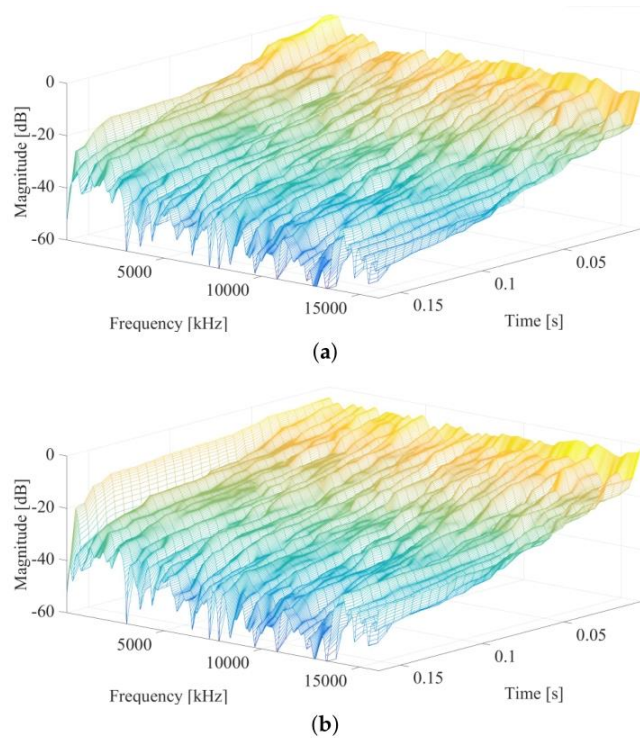


Рис. 13 Значення EDR: (a) без вирівнювання, (b) після вирівнювання

Рисунок 13 демонструє значення EDR розраховане до та після процедури вирівнювання. За часовими координатами на графіки демонструють зменшення часу затухання, тоді як за частотними координатами можна спостерігати зменшення частотних піків. Загалом, після процедури вирівнювання було отримано більш узагальнену залежність, зі зменшенням (згладжуванням) піків та прогалин.

9.1.4 Акустичні параметри

Якість акустичного сигналу може бути оцінена через використання об'єктивних параметрів оцінки, що базуються на імпульсній характеристиці (імпульсному відгуку) кімнати.

Акустичні параметри, отримані в результаті застосування об'єктивних параметрів були вперше застосовані для оцінки вирівнювання імпульсної

характеристики (імпульсного відгуку) (RRE) кімнати в джерелі [6]. Наступні акустичні параметри були використані в більшості статей за тематикою RRE:

- Коефіцієнт визначеності, що визначається як процентне відношення енергії перших 50 або 80 мілісекунд імпульсної характеристики після основного піку до решти енергії імпульсної характеристики кімнати (D50 або D80).

- Коефіцієнт розбірливості, що визначається як логарифмічне відношення енергії перших 50 або 80 мілісекунд імпульсної характеристики після основного піку до решти енергії імпульсної характеристики кімнати (C50 або C80).

- Ранній час затухання, що визначається як час за який відбудеться початкове затухання сигналу на 10 дБ, помножений на 6 (EDT).

- Коефіцієнт відношення сигналу до реверберацій (DRR) яке ще називають відношення енергії сигналу до енергії реверберацій, що визначається як відношення між основним піком та всією іншою імпульсною характеристикою (імпульсним відгуком) кімнати.

- Центральний час, що є «центром гравітації» енергії імпульсної характеристики кімнати.

9.2 Суб'єктивне оцінювання

Для правильної оцінки якості звуку, слухове тестування необхідно проводити відповідно до стандартизованої процедури. Велика кількість методик та процедур суб'єктивної оцінки акустичних систем може бути знайдена в літературі. Проте, фокусуючись на параметрі вирівнювання імпульсної характеристики (імпульсного відгуку) (RRE) кімнати та відштовхуючись від сучасного рівня та стану досліджень за тематикою, суб'єктивна оцінка RRE має відповідати наступним стандартам:

ITU-R BS.1116-1 – «Методи суб'єктивної оцінки малих недоліків в аудіо системах, зокрема багатоканальних аудіо систем»;

ITU-R BS.1534-1 – «Методи суб'єктивної оцінки проміжного рівня якості систем кодування»;

ITU-R BS.1284-1 – «Загальні методи суб'єктивної оцінки якості звуку.

Всі ці стандарти містять у собі методології тестування, процедури тестування та статистичні методи обробки отриманих результатів. Проте, через широту даної тематики, обговорення буде сфокусоване лише на найбільш актуальних процедурах, що застосовні до вирівнювання імпульсної характеристики (RRE).

Стандарт ITU-R BS.1284-1 надає директиви щодо загальної оцінки якості сприйнятого звуку, та був застосований в джерелах для оцінки RRE. Варто помітити, що стандарт ITU-R BS.1284-1 базується на стандарті ITU-R BS.1116-1. Відповідно до директив розміщених в джерелі, при виборі учасників експерименту перевага має надаватися досвідченим слухачам для «отримання кращого та швидшого визначення найбільш ймовірних результатів в довгостроковій перспективі». Процес суб'єктивного оцінювання побудований як порівняльний тест, де учасники мають надати оцінку використовуючи семибальну шкалу з рекомендованим градуванням у 1 бал, як показано в джерелі. Тест побудований на попарному порівнянні з посиленнями (оригіналами, незміненими зразками), оцінки виставляються після повторюваного прослуховування (4 рази послідовно) визначеної програмної послідовності зразків. У випадку оцінювання процедури вирівнювання, була використана наступна послідовність:

- Незмінена послідовність без вирівнювання;
- Та ж сама послідовність, вирівняна одним з обраних методів вирівнювання;
- Повторення незміненої послідовності без вирівнювання;
- Повторення тієї ж самої послідовності, вирівняної одним з обраних методів вирівнювання

Як було рекомендовано в (244), протяжність зразків не має перевищувати 20 секунд, тому протяжність була обмежена на рівні 15-20 секунд. Також, були прийняті міри до того, прослуховувані зразки не мали розривів. Для кожного зразка, порядок демонстрації різних методів вирівнювання був обраний випадковим чином (рандомізований), та слухачу не було відомо який метод

вирівнювання наразі тестується. Відповідно до рекомендацій, перед тестуванням слухачам було проведено тренування. Як вказано в джерелі №145, для ознайомлення з процедурою тестування, тестовими зразками та тестовим середовищем, слухачі мали нагоду прослухати окремо кожний аудіо зразок з впливом кожного з методів вирівнювання що оцінюються. Тоді як стандарт ITU-R BS.1284-1 передбачає декілька критеріїв для суб'єктивної оцінки якості звуку, в роботах №75 та 145 були використані три критерії: «чіткість» (всі фактичні деталі зразку зрозумілі), «тембр, забарвленість» (можна навести точний опис прослуханого зразка), «загальне враження» (узагальнення всього звукового «портрету» та міжзв'язків попередніх параметрів). Для оцінки вирівнювача імпульсної характеристики кімнати з використанням зразків з різною спектральною характеристикою (спектральним складом) використовувалися зразки, що відносяться до різних музичних жанрів. Зрештою, отримані результати були оброблені для отримання усереднених значень та довірчих інтервалів. Значення похибки в 0.05 було отримане при визначенні довірчих інтервалів.

10. Питання, що набирають популярність та нові тенденції

10.1 Особисті звукові зони

В останні роки зростає увага до можливості відтворення різного контенту в суміжних, просторово обмежених зонах для декількох слухачів через зменшення завад між зонами. Ці розробки відомі як персональні звукові зони, мультizonний синтез або мультizonний звуковий контроль. Нещодавній огляд за даною тематикою може бути знайдений в (246), а більш детальний опис в ряді статей за тематикою (247-274). На сьогоднішній день, досягнє придушення завад між зонами обмежене різними акустичними та практичними обмеженнями, що обмежує сферу практичного застосування технології. В (262) наведено огляд головних ускладнень при реалізації мультizonного звукового контролю в середовищах з відбиттям (в реверберуючих середовищах). Забезпечення придушення завад та компенсація властивостей кімнати, стійкого до змін та невизначеностей залишається основними проблемами, що потребують

розв'язання. Підхід до вирівнювання характеристик кімнати через (для) контроль звукового тиску визначеної області простору в поєднанні з переставленням звукового поля в хвильовій області наведено в (262). Підхід визначено стійким в низькочастотній області, проте неефективним на високих частотах, де звукове поле відбитого сигналу (ревербераційне звукове поле) розсіюється, що призводить до необхідності застосування великої кількості динаміків.

10.2 Портативні пристрої

В останні роки, застосування портативних пристроїв зросло неймовірно, дозволивши їм досягти високого ступеня розповсюдженості. Проте, через характеристики використовуваних динаміків та їх взаємодію з внутрішнім середовищем кімнати, велика кількість з них здатна задовольнити лише базові потреби з забезпечення якісного звучання. Ця ситуація може бути частково виправлена шляхом вивчення акустичних характеристик даних пристроїв та застосуванні просунутих акустичних технологій для їх покращення. В (275), представлено застосування процедури багатоточкового вирівнювання для покращення неідеального відгуку (імпульсної характеристики) портативної звукової системи, такої як мобільний телефон. Результати об'єктивних вимірювань та суб'єктивного прослуховування довели позитивний ефект застосування алгоритмів на звучання персональних портативних пристроїв. В (276), продемонстровано застосування статичного та адаптивного алгоритму для лінеаризації частотного відгуку (характеристики) до звукових систем портативних комп'ютерів. Суб'єктивні прослуховування тестових записів підкреслюють покращення в сприйнятті звукової інформації слухачем, підтверджуючи корисність даного підходу.

10.3 Нелінійна еквалізація (вирівнювання нелінійності)

Системи відтворення звуку можуть піддаватися небажаному впливу не лише з боку акустичного впливу кімнати, а й впливу нелінійності характеристик підсилювача та динаміків, що входять до їх складу. Для виключення даного впливу, в джерелах (277-279) пропонується застосування вирівнювачів

(еквалайзерів), що містять фільтри Вольтерра в своїй будові, до моделі підсилювач-динамік-екран до передачі вихідного сигналу на динаміки. Таким чином стає можливим вирівнювання не лише лінійної, але й нелінійної характеристики системи, через використання адаптивних методів.

10.4 Вирівнювання характеристик кімнати з використанням рухомого мікрофону

Одним з головних недоліків застосування багато-точкового вирівнювання є необхідність вимірювання імпульсного відгуку (характеристики) кімнати (RIR), що потребує великої кількості часу для отримання чіткої просторової розподільчої здатності всередині зони прослуховування. Вирішенням цієї проблеми може стати застосування розподіленої в часі системи ідентифікації. В ній імпульсні відгуки (характеристики) кімнати (RIR) вимірюються з застосуванням динамічного методу, що базується на використанні єдиного мікрофону що рухається кімнатою, на відміну від методу з одночасним незалежним вимірюванням імпульсних відгуків (характеристик) кімнати (RIR). Ця процедура дозволяє отримати щільну сітку імпульсних відгуків (характеристик) кімнати (RIR) з одного просторово неперервного вимірювання, що може бути використане в процесі багато-точкового вирівнювання для визначення функції-оригіналу та вирівнювального фільтрів.

11. Висновки

В цій роботі було досліджено сучасний стан вивчення та історичний шлях досліджуваної тематики вирівнювання імпульсної характеристики кімнати. Для підкреслення та демонстрації еволюційного розвитку та потенціалу дослідження тематики, було наведено різні класифікації існуючих рішень. Перша класифікація заснована на визначенні кількості імпульсних відгуків, що були враховані при створенні вирівнювального фільтру (наприклад одно-точкові та багато-точкові вирівнювачі(еквалайзери)). Перший ефективний лише в обмеженій зоні навколо точки вимірювання, тоді як другий здатний до розширення зони вирівнювання та роботи з розрізненими імпульсними відгуками різних зон кімнати. Друга класифікація проводиться виходячи з

методу вимірювання імпульсних характеристик: миттєвого або неперервного (статичні та динамічні (або адаптивні) підходи). Перший заснований на апріорному вимірюванні імпульсних відгуків, тоді як другий базується на постійному оновленні імпульсних характеристик і таким чином дозволяє вирівнювачу (еквалайзеру) адаптуватися до тимчасових змін в середовищі вимірювання. Всередині даної загальної класифікації ми маємо виділити підходи з перед-обробкою, що використовуються для виділення завад(шумів) звукового діапазону частот викликаних помилками вирівнювання, що з'явилися через зміни в імпульсній характеристиці кімнати, підходи з виділенням мінімальної та змішаної фази, прямі та непрямі підходи, що застосовуються у різних методах побудови вирівнювачів (еквалайзерів), та фільтри хвильової області для вирівнювання звуку великих багатоканальних систем відтворення. Відповідно до даної класифікації, в роботі були описані різні підходи до вирішення даної задачі. В Таблиці 1 зведено сучасні методи вирішення задачі, відсортовані відповідно до класифікації наведеної вище, наприклад підходи з перед-обробкою, підходи з виділенням мінімальної та змішаної фази, статичні та динамічні (або адаптивні) підходи, приклад одно-точкові та багато-точкові підходи, прямі та непрямі методи, базуючись на визначенні даному у розділі 4.6, та фільтри хвильової області. Вона доводить, що один метод може підпадати під декілька пунктів класифікації, розширюючи потенціал та ефективність методології (мається на увазі загалом тематика вирівнювання імпульсної характеристики кімнати). В даному контексті, інструментальні вимірювання та суб'єктивні оцінювання вирівнювання стають критичними: були розглянуті деякі приклади сучасних підходів (методологій проведення досліджень) у даній тематиці. На закінчення, було наведене загальне обговорення зростаючих методологій та нових тенденцій в галузі вирівнювання імпульсної характеристики кімнати. З нього можна зробити висновок, що зростаюча доступність приватних засобів прослуховування призведе до більш частого застосування технологій в галузі вирівнювання імпульсної характеристики кімнати з метою підвищення їх продуктивності, покращення їх характеристик.

Література до лекції 2.2

1. Stefania Cecchi, Alberto Carini and Sascha Spors. Room Response Equalization—A Review. Appl. Sci. 2018, 8, 16; doi:10.3390/app8010016. - P.1-47.

Розділ 3. Завдання до МКР та самостійної роботи

Розділ 3.1. Завдання із МКР

1. **Термін виконання** – а) результати прослуховування у вигляді електронних файлів – **через 2 тижні** від моменту отримання завдання; б) письмовий звіт із постановкою завдання, отриманими результатами у графічній формі, коментарями то отриманих результатів – **через 4 тижні** від моменту отримання завдання.
2. **Вихідні дані** до МКР: в архівних файлах **Project_2020_bin_209.rar** та **Project_2020_bin_438.rar**
3. **Після розархівування** буде отримано:
 - папку **Programs** із допоміжними програмами та файлами;
 - файл **! Readme_209.txt**;
 - файл запуску тестування **begin_testing_209.m**;
 - файл обробки результатів тестування **rzlts_processing_bin_209.m**.
4. **Що треба зробити**:
 - Запустити **begin_testing_209.m** та виконати прослуховування
 - Запустити **rzlts_processing_bin_209.m** - з'явиться файл **mean_results.mat** (результати) та 2 графіки
 - Зберегти отримані графіки у форматі **png** та розмістити їх у папці **Results209_Lastname**
 - Зархівувати папку **Results209_Lastname** та переслати викладачу.

Розділ 3.1. Завдання до самостійної роботи

Самостійна робота полягає у роботі із оригінальною англomовною статтею *Stefania Cecchi, Alberto Carini and Sascha Spors. Room Response Equalization—A Review. Appl. Sci. 2018, 8, 16; doi:10.3390/app8010016*. - P.1-47 та її перекладом (розділ 2.2 даного конспекту).

Потрібно:

- 1) прочитати статтю та виробити рекомендації із поліпшення перекладу;

- 2) сформулювати та викласти свою точку зору на актуальність проблеми, що розглянуто в огляді;
- 3) сформулювати та обґрунтувати свою точку зору на найбільш перспективні шляхи рішення задачі, розглянутої в огляді.

Обговорення зазначених питань виконується на спеціальному семінарі, при цьому активність обговорення оцінюється за 5-бальною системою.

Базова література

- 47.Продеус А., Дидковский В., Дидковская М. Акустическая экспертиза и коррекция коммуникационных каналов. Монография. LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, Saarbrücken, Deutschland, 2017, ISBN: 978-3-330-04591-0
- 48.Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации. Монография. – К., «Имекс-ЛТД», 2008. – 420 с.
- 49.Ю.С. Костючок, Л.С. Мартинович, Д.Е. Моторнюк, В.А. Нечитайло, А.В. Храпачевский, А.Н. Продеус. Акустическая паспортизация учебных помещений. - ISSN 1811-4512. Electronics and Communications, 2016, Vol.21, No.2(91). – P.63-70
- 50.О. Дворник, А. Продеус, М. Дідковська, Д. Моторнюк, "Апаратно-програмний комплекс «Штучна голова». Частина 1. Коригування частотної характеристики тракту," ISSN 2523-4447, Мікросистеми, електроніка та акустика, 2020, vol. 22, no. 1, сс. 56-64. DOI: <https://doi.org/10.20535/2523-4455.mea.198431>
- 51.О. Дворник, А. Продеус, М. Дідковська, Д. Моторнюк, "Апаратно-програмний комплекс «Штучна голова». Частина 2. Оцінювання розбірливості мови в аудиторіях," Microsystems, Electronics and Acoustics, vol. 22, no. 3, 2020, pp. 48-55. DOI: <https://doi.org/10.20535/2523-4455.mea.209928>
- 52.Продеус А.М. Комп'ютерна обробка акустичних сигналів [Електронний ресурс]. Навчальний посібник. Київ: КПП, 2018, 237 с. <https://ela.kpi.ua/handle/123456789/25161>
- 53.Prodeus, A., Bukhta, K., Morozko, P., Serhiienko, O., Kotvytskyi, I., Shherbenko I. (2018). Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility. IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO`2018), Kyiv, Ukraine, 533-538. doi: <https://doi.org/10.1109/ELNANO.2018.8477568>

- 54.Сергиенко А.Б. Цифровая обработка сигналов. СПб: Питер., 2002 г., 606 с.
- 55.Богнер Р., Константи́нидис А. (ред.) Введение в цифровую фильтрацию. М.: Мир, 1976, 216 с.
- 56.Leccese, F., Rocca, M., Salvadori, G. (2018) Fast estimation of Speech Transmission Index using the Reverberation Time: Comparison between predictive equations for educational rooms of different sizes. Applied Acoustics, 140, 143–149. doi: <https://doi.org/10.1016/j.apacoust.2018.05.019>
- 57.Nocke, C. (2016). New standards in Architectural Acoustics - a German view. ACOUSTICS`2016, Brisbane, Australia. Available in: https://www.acoustics.asn.au/conference_proceedings/AASNZ2016/papers/p114.pdf
- 58.DIN 18041:2004-05. (2004). Hörsamkeit von kleinen und mittleren Räumen. (Acoustical quality in small to medium-sized rooms.),” Beuth Verlag GmbH. Available in: <https://www.beuth.de/de/norm/din-18041/69640391>
- 59.ANSI/ASA S12.60. (2015). Acoustical performance criteria, design requirements, and guidelines for schools. Part 1: permanent schools. American National Standards Institute, Inc. (ANSI) and Acoustical Society of America (ASA) New York, USA. Available in: <https://webstore.ansi.org/standards/asa/ansiasas12602010partr2015>
- 60.International standard ISO 3382-1. (2009). Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces. First edition. Available at: <https://www.iso.org/obp/ui/#iso:std:iso:3382:-1:ed-1:v1:en>
- 61.British standard BS EN 60268-16:2011. (2012). Sound systems equipment - Part 16: Objective rating of speech intelligibility by speech transmission index. BSI. Available at: https://infostore.saiglobal.com/en-us/Standards/BS-EN-60268-16-2011-217511_SAIG_BSI_BSI_512663/
- 62.Acoustic design of schools: performance standards. (2015). Building bulletin 93. UK Department for Education, UK Education Funding Agency. Available at: <https://www.bloconconsulting.co.uk/building-bulletin-93-bb93/>

- 63.Canning, D., James, A. (2012). The Essex Study. Optimised classroom acoustics for all. The Association of Noise Consultants. Available at: https://www.ndcs.org.uk/media/1808/the_essex_study_optimised_classroom_acoustics_for_all1.pdf
- 64.IEC 60268-16. (2016). Sound system equipment – Part 16: objective rating of speech intelligibility by speech transmission index. International Electrotechnical Commission (IEC), Geneva, Switzerland. Available at: https://infostore.saiglobal.com/en-us/Standards/IEC-60268-16-4-0-569917_SAIG_IEC_IEC_1301856/

Додаткова література

- 65.Дідковський В.С., Дідковська М.В., Продеус А.М. Комп'ютерна обробка акустичних сигналів. Навчальний посібник. – К., «Імекс-ЛТД», 2010. – 420 с.
- 66.Стандарти та вимоги до акустики шкільних класів та аудиторій вишів Америки, Великобританії, Німеччини, Росії, України. Доступно в системі дистанційного навчання НТУУ «КПІ ім. І. Сікорського» та в Google Classroom.
- 67.M. Jeub, M. Schafer, and P. Vary, “A binaural room impulse response database for the evaluation of dereverberation algorithms,” in Proc. Int. Conference on Digital Signal Processing (DSP), Santorini, Greece, 2009.
- 68.Eggenschwiler, K. (2005). Lecture Halls - Room Acoustics and Sound Reinforcement. Forum Acusticum, Budapest, 1-7. Available at: https://www.researchgate.net/publication/238114069_Lecture_Halls_-_Room_Acoustics_and_Sound_Reinforcement
- 69.Bradley, J., Reich, R., Norcross, S. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. Journal of the Acoustical Society of America, 106 (4), Pt. 1. doi: <https://doi.org/10.1121/1.427932>
- 70.Bradley, J., Sato, H., Picard, M. (2003). On the importance of early reflections for speech in rooms. Journal of the Acoustical Society of America, 113 (6), 3233-3244. doi: <http://dx.doi.org/10.1121/1.1570439>
- 71.Hu, Y., Kokkinakis, K. (2013). Effects of early and late reflections on intelligibility of reverberated speech by cochlear implant listeners. Journal of the Acoustical Society of America, 135 (1). doi: <https://doi.org/10.1121/1.4834455>
- 72.Yang, W., Bradley, J. (2009). Effects of room acoustics on the intelligibility of speech in classrooms. Journal of the Acoustical Society of America, 125 (2), 1-12. doi: <https://doi.org/10.1121/1.3058900>

73. Bradley, J. (2011). Review of objective room acoustics measures and future needs. *Applied Acoustics*, 72 (10), 713-720. doi: <https://doi.org/10.1016/j.apacoust.2011.04.004>
74. Arweiler, I., Buchholz, J., Dau T. (2009). Speech intelligibility enhancement by early reflections. ISAAR 2009: Binaural Processing and Spatial Hearing, 2nd International Symposium on Auditory and Audiological Research, Elsinore, Denmark. Available at: <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-29>
75. Sari, L., Utami S., Sarwono, J. (2018). Objective and subjective acoustics measurement of audience seating areas in a medium size auditorium. *Journal of Physics: Conf. Series* 1075. doi: 10.1088/1742-6596/1075/1/012015
76. Choi, Y.-J. (2020). The intelligibility of speech in university classrooms during lectures. *Applied Acoustics*, 162, 107211. doi: <https://doi.org/10.1016/j.apacoust.2020.107211>
77. Eldakdoky, S. (2017). Optimizing acoustic conditions for two lecture rooms in Faculty of Agriculture, Cairo University. *Ain Shams Engineering Journal*, 8, 481–490. doi: <http://dx.doi.org/10.1016/j.asej.2016.08.013>
78. Prodeus, A., Didkovska, M., Motorniuk, D., Dvornyk, O. (2020). The Effects of Noise, Early and Late Reflections on Speech Intelligibility. *IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO`2020)*, Kyiv, Ukraine, 488–492. doi: <https://doi.org/10.1109/ELNANO50318.2020.9088854>
79. Prodeus, A., Didkovska, M. (2020). Objective assessment of speech intelligibility in small and medium-sized classrooms. *IEEE Int. Scientific-Practical Conf. on Problems of Infocommunications, Science and Technology (PIC S&T`2020)*, Kharkiv, Ukraine. Available at: https://www.researchgate.net/publication/347490796_Objective_Assessment_of_Speech_Intelligibility_in_Small_and_Medium-Sized_Classrooms
80. Prodeus, A., Didkovska, M., Kukharicheva, K., Motorniuk, D. (2020). Modeling the Influence of Early Sound Reflections on Speech Intelligibility. *IEEE 6th International Conference on Methods and Systems of Navigation and*

- Motion Control (MSNMC`2020), Kyiv, Ukraine, 47-50. Available at: https://www.researchgate.net/publication/346080603_Modeling_the_Influence_of_Early_Sound_Reflections_on_Speech_Intelligibility
81. Jacob, K. (1988). Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms. 85th Convention of the Audio Engineering Society, Los Angeles, 1020-1030. Available at: <https://secure.aes.org/forum/pubs/journal/?elib=6056>
 82. Lochner J., Burger, J. (1964). The influence of reflections on auditorium acoustics,” J. Sound Vib. 1, 426–454. doi: [https://doi.org/10.1016/0022-460X\(64\)90057-4](https://doi.org/10.1016/0022-460X(64)90057-4)
 83. Bistafa, S., Bradley, J. (2000). Reverberation time and maximum background noise level for classrooms from a comparative study of speech intelligibility metrics. J. Acoust. Soc. Am., 107, 861–875. doi: <https://doi.org/10.1121/1.428268>
 84. Gold, B., Morgan, N., Ellis, D. (2011). Speech and audio signal processing. Processing and Perception of Speech and Music. John Wiley & Sons, Inc., 679. Available at: <https://ieeexplore.ieee.org/book/8039862>
 85. Nestoras, C., Dance, S. (2013). The Interrelationship Between Room Acoustics Parameters as Measured in University Classrooms Using Four Source Configurations. Building acoustics, 20 (1), 43–54. doi: <https://doi.org/10.1260/1351-010X.20.1.43>
 86. Aachen Impulse Response Database. Available at: <https://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/>
 87. Duran, S., Ausiello, L., Battaner-Moro, J. (2019). Acoustic Design Criteria for Higher-Education Learning Environments. Proceedings of the Institute of Acoustics, 41 (3), 1-12. Available at: <https://pure.solent.ac.uk/en/publications/acoustic-design-criteria-for-higher-education-learning-environmen>

88. Jedidi, M., Boulila, A. (2016). Acoustic study of an auditorium by the determination of reverberation time and speech transmission index. *Int. J. Architect. Eng. Urban Plan*, 26(1), 25-32. Available at: <http://ijaup.iust.ac.ir/article-1-312-en.pdf>
89. Silva, L., Santos, R. (2013). Acoustical Comfort in Primary School Classrooms in the City of Joao Pessoa, Paraiba, Brazil. *Journal of Ergonomics*, S:1. doi: <https://doi.org/10.4172/2165-7556.S1-001>
90. Renniesa, J., Schepker, H., Holube, I., Kollmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J. Acoust. Soc. Am.*, 136 (5), 2642–2653. doi: <https://doi.org/10.1121/1.4897398>
91. Kuttruff, H. (2009). *Room Acoustics*. Fifth edition. Spon Press 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN. Available at: <https://www.amazon.com/Room-Acoustics-Fifth-Heinrich-Kuttruff/dp/0415480213>
92. Steeneken, H. (2014). Forty years of speech intelligibility assessment (and some history). *Proceedings of the Institute of Acoustics*, 36, Pt.3. Available at: <http://www.steeneken.nl/wp-content/uploads/2014/04/IOA-Forty-years-keynote.pdf>

Додатки

Додаток до підрозділу 1.7. Розробка пакетів програм для суб'єктивного оцінювання розбірливості мови в приміщеннях 209 та 438

Вступні зауваження

Будемо використовувати матеріал, що міститься в 2-х папках:

- 1) D:\A_Raboty_3\Kpi_new\Acoustic\E_Comput_acoust_syst\2020_КАС_учеба\2_Lect
t
- 2) D:\A_Raboty_3\Kpi_new\Acoustic\!Проект бін.
слух\Project_2019_bin_3\Project_bin_3

Далі ці папки будемо для простоти нумерувати як 1-1 та 1-2.

Нові пакети програм будемо розміщати в папках:

- 1) D:\A_Raboty_3\Kpi_new\Acoustic\!Проект бін. слух\Project_2020_bin_209
- 2) D:\A_Raboty_3\Kpi_new\Acoustic\!Проект бін. слух\Project_2020_bin_438

Далі ці папки будемо для простоти нумерувати як 2-1 та 2-2.

Формування mat-файлів із IX приміщень

Програму формування mat-файлів із IX приміщень розміщено в Додатку 1 та в папці 1-1.

Вхідними даними для неї слугують попередньо обчислені wav-файли IX приміщення 209 (6 файлів типу **RIR_209_Np.wav**, де N – номер точки) та приміщення 438 (5 файлів типу **RIR_438_Np_opyt2.wav**, де N – номер точки).

Результатами її роботи є файли **rirs209_table_LR.mat** та **rirs438_table_LR.mat**, які також розміщено в папці 1-1.

Формування пакету програм для приміщення 209

Таке формування виконуємо покроково.

Крок 1. Копіюємо із папки 1-2 в папку 2-1:

- 1) папку **Programs**;
- 2) файли **channel_testing.wav**, **melodiya.wav**, **rzlts_processing_bin.m**,
table_reading_assist_bin_50ms.m.

Крок 2. Тут та далі працюємо в папці 2-1:

- 1) в папку **Programs** вставляємо **rirs209_table_LR.mat** замість **rirs_table_LR.mat**;
- 2) **wav**-файли переміщуємо в папку **Programs**;
- 3) файл **table_reading_assist_bin_50ms.m** перейменовуємо в **begin_testing_209.m**;

Крок 3. Робимо правки в скриптах (в Matlab'і):

- 1) в **begin_testing_209.m** – правимо: вступні коментарі та паузи між ними; адреси музичних файлів; коментар до налагоджування навушників; значення параметрів SNR0 та T60D; значення `params.cnt4 == 90` на `params.cnt4 == 24` (24 – це 8 сполучень IX та SNR, помножених на 3 таблиці); щоб надалі не мучитися із цим, автоматизуємо обчислення кількості таблиць, що слухаються (`num_SNRT = length(SNRT)`; `num_T60D = length(T60D)`; `num_tables = num_SNRT*num_T60D*3`; `if params.cnt4 == numtables`); **[params] = assistant_bin_3(kat_results, NT,num_tables,num_SNRT,num_T60D)**; `kat_results = ['ResultsTraining209_' lastname]`; `kat_results = ['Results209_' lastname]`;
- 2) створюємо **assistant_bin_3(kat_results, NT,num_tables,num_SNRT,num_T60D)**, де правимо: `maxnfiles4 = num_tables`; `for nsnr = 1:num_SNRT`; `for nt60 = 1:num_T60D`; `params.Fs = 44100`; `rirs_LR = load('rirs209_table_LR.mat')`;
- 3) в **distorSystem_bin_50ms** – перейменовано на **distorSystem_bin_50ms_cor**; правимо: `detVoice = vad_LP(inputSignalFs,Fs)` (виправлено - було `fs`, що невірно, оскільки перед відправкою в `vad_LP` мовний сигнал передискретизовано із `fs` в `Fs`);
- 4) в **begin_testing_209.m** – правимо із врахуванням правок п.3: замість **distorSystem_bin_50ms** пишемо **distorSystem_bin_50ms_cor**.
- 5) **rzlts_processing_bin.m** перейменовуємо в **rzlts_processing_bin_209.m** та правимо: значення параметрів SNR0 та T60D; додаємо **num_SNRT = length(SNRT)** та **num_T60D = length(T60D)**; враховуємо можливість обробки результатів, отриманих в режимі Тренування; робимо `val4 = zeros(num_T60D,num_SNRT,3)` та міняємо границі циклів збирання даних в 3-мірну матрицю; побудова графіків результатів – коригуємо графіки до 4-х точок приміщення та будуємо 2 графіки – CVC(SNR) та CVC(D); збереження результатів усереднення в файлі-структурі – зменшуємо кількість графіків до 4-х точок приміщення;
- 6) в **begin_testing_209.m** – правимо: видаляємо пропозицію виправити попереднє слово;

Перевірка роботи для приміщення 209

Проведено оцінювання, в режимі ТРЕНУВАННЯ, розбірливості для 2-х випадків: SNR0 = -5 дБ та SNR0 = 5 дБ.

Що добре: комплекс програм показав себе як працездатний.

Недолік: залежність CVC(SNR) виявилася надто примітивною, оскільки графік будувався за 2 крайніми точками. Крім того, у випадку SNR = 5 дБ розбірливість виявилася практично 100%.

Якщо додати ще один варіант SNR, тоді в режимі РОБОТА час тестування буде 3 години замість 2 годин, що теж не прийнятно.

Тому залишаємо 2 значення SNR, але робимо їх -6 дБ та 0 дБ.

Формування пакету програм для приміщення 438

- 1) копіюємо вміст папки 2-1 в папку 2-2; наступні зміни робимо лише у папці 2-2;
- 2) в папку **Programs** вставляємо **rirs438_table_LR.mat** замість **rirs209_table_LR.mat**;
- 3) у файлі **begin_testing_209.m** робимо: `kat_results = ['ResultsTraining438_' lastname]; kat_results = ['Results438_' lastname];` робимо заміну: **assistant_bin_3.m** замінюємо на **assistant_bin_4.m**;
- 4) у файлі **assistant_bin_4.m** робимо заміну: `rirs_LR = load('rirs438_table_LR.mat');`
- 5) у файлі **rzlts_processing_bin_438.m** робимо правки: `T60D = [2.4 4.8 6.9];` `kat_results = ['ResultsTraining438_' lastname]; kat_results = ['Results438_' lastname];` також правимо команди побудови графіків та збереження результатів у mat-файлі.

Перевірка роботи для приміщення 438

Проведено оцінювання, в режимі ТРЕНУВАННЯ, розбірливості для 2-х випадків: SNR0 = -5 дБ та SNR0 = 5 дБ. Комплекс програм показав себе як працездатний.

Тексти програм

Скрипти програм для к.209 наведені в Додатках. Для кімнати 438 скрипти аналогічні. Робочі пакети програм розташовано за адресами:

D:\A_Raboty_3\Kpi_new\Acoustic\!Проект бін. слух\Project_2020_bin_209

D:\A_Raboty_3\Kpi_new\Acoustic\!Проект бін. слух\Project_2020_bin_438

```

% === make_rirs209_438_matfile.m ===
% програма створення mat-файлів із записами IX 209 та 438 кімнат
%
% 1) вказуємо номер кімнати
% room_num = 209;
room_num = 438;

% 2) прочитуємо потрібні IX;
% формуємо mat-файл для кімнати 209:
% save('rirs209_table_LR.mat','r1','r2','r3','r4','r5','r6');
if room_num == 209
    [r1,Fs] = audioread(['RIR_' num2str(room_num) '_1p.wav']);
    r2 = audioread(['RIR_' num2str(room_num) '_2p.wav']);
    r3 = audioread(['RIR_' num2str(room_num) '_3p.wav']);
    r4 = audioread(['RIR_' num2str(room_num) '_4p.wav']);
    r5 = audioread(['RIR_' num2str(room_num) '_5p.wav']);
    r6 = audioread(['RIR_' num2str(room_num) '_6p.wav']);
    save('rirs209_table_LR.mat','r1','r2','r3','r4','r5','r6','Fs');
end

% 3) аналогічно формуємо mat-файл для кімнати 438:
% save('rirs438_table_LR.mat','r1','r2','r3','r4','r5');
if room_num == 438
    [r1,Fs] = audioread(['RIR_' num2str(room_num) '_1p_opyt2.wav']);
    r2 = audioread(['RIR_' num2str(room_num) '_2p_opyt2.wav']);
    r3 = audioread(['RIR_' num2str(room_num) '_3p_opyt2.wav']);
    r4 = audioread(['RIR_' num2str(room_num) '_4p_opyt2.wav']);
    r5 = audioread(['RIR_' num2str(room_num) '_5p_opyt2.wav']);
    save('rirs438_table_LR.mat','r1','r2','r3','r4','r5','Fs');
end

```

Додаток 2

Програма begin_testing_209.m

```

%% === begin_testing_209.m === слухання таблиць ===
%
% Мета даної програми - автоматизація оцінювання розбірливості
% звукосполучень за результатами прослуховування 3-х наборів (таблиць)
% таких звукосполучень. Завдяки підпрограмі-функції assistant_bin.m, яка
% контролює повноту виконання поставленого завдання, слухачу не треба
% пам'ятати, що вже зроблено і що ще залишилося зробити...
%
% Розробка програми: А.Продеус - 2017 р. Модернізація програми: О.Чайка,
% М.Ярошенко, А.Продеус - 2017-2019 рр. Модернізація полягала в наступному:
% 1) випадковий порядок прослуховування - М.Ярошенко; 2) можливість
% повернутися на крок назад - О.Чайка; 3) використання 3-ї еталонної
% таблиці та остаточна корекція програми - А.Продеус; 4) налаштування
% програми до проекту - "Суб'єктивне оцінювання розбірливості мовлення із
% використанням бінауральних імпульсних характеристик приміщення" -
% А.Продеус
%
% Структура програми:
%
% 1. Запит вхідних даних для програми assistant_bin_2.m:

```

```

% 1.1. Вибір режиму роботи
% 1.2. Запит прізвища виконавця
% 1.3. Запит номерів 3-х таблиць
% 2. Запуск програми assistant_bin_2.m
% 3. Завантаження wav-таблиці із потрібним номером
% 4. Початок циклу прослуховування сигналів із wav-таблиці
% 5. Підрахунок частки правильно сприйнятих звукосполучень
% 6. Запит про подовження слухання таблиць
%
% Перелік дослідів та підрахунок часу їх виконання:
%-----
% Кіль-   |
% кість   |           Сутність дослідів
% дослідів |
%-----
%      8   | Спотворення БІЛИМ шумом із snr0 = -6 дБ та 0 дБ
%          | та реверберацією для 4-х точок приміщення 209
%
% В кожному досліді слухається 3 таблиці звукосполучень,
% тобто загальна кількість таблиць, що прослуховуються, складає 24.
% Оскільки прослуховування однієї таблиці потребує приблизно 4-5 хвилин,
% маємо 24x5 = 120 хвилин або 2 години.
%
% Copyright @ А.Продеус,О.Чайка, М.Ярошенко, 2017-2020 рр.

%% 1. Вибір режиму роботи:

clc      % очистка командного окна
clear    % очистка рабочего пространства (workspace)

[x,Fs] = audioread('Programs\melodiya.wav'); % Fs = 44100 Hz
playerObj1 = audioplayer(x,Fs);
play(playerObj1);
pause(length(x)/Fs + 2.5);

disp('Перевірте правильність налаштування навушників (канали, гучність');
disp('Якщо не все гаразд, тоді ПЕРЕД натисканням 0 поправте канали та гучність');
[y,Fs] = audioread('Programs\channel_testing.wav'); % Fs = 44100 Hz
playerObj2 = audioplayer(y,Fs);
zvuk = 0;
while zvuk == 0
    play(playerObj2);
    pause(length(y)/Fs);
    disp(' ');
    zvuk = input('Звукові канали та гучність в нормі? Так - 1, Ні - 0 ');
end

clc

disp('В режимі ТРЕНУВАННЯ прослуховується 3 звукосполучення кожної із таблиць');
disp('В режимі РОБОТА прослуховуються всі 50 звукосполучень кожної із таблиць');
disp(' ');
wt = input('Оберіть режим роботи? 1 - РОБОТА; 0 - ТРЕНУВАННЯ: '); %!!!
if wt == 0
    %numtren = input('Тренування: задайте кількість звукосполучень (від 3 до 5): ');
    nums = 3; % кількість звукосполучень в режимі ТРЕНУВАННЯ
else

```

```

    nums = 50;    % кількість звукосполучень в режимі РОБОТА
end

%% 2. Запит вхідних даних для assistant_bin.m (прізвище, номери таблиць)
lastname = input('Вкажіть ваше прізвище латиною (без лапок): ', 's');
disp('Вкажіть в порядку зростання номери трьох таблиць для слухання :');
NT = zeros(1,3);
NT(1) = input('Вкажіть номер першої таблиці: N1 = '); %!!!
NT(2) = input('Вкажіть номер другої таблиці: N2 = '); %!!!
NT(3) = input('Вкажіть номер третьої таблиці: N3 = '); %!!!
SNRT = [-6 0];    % таблиця значень потрібних SNR0
num_SNRT = length(SNRT);
T60D = [3.5 5.2 8.6 10.3]; % дистанції гучномовець-мікрофон
num_T60D = length(T60D);
num_tables = num_SNRT*num_T60D*3;

clc

disp('Переходимо до прослуховування звукосполучень типу "гуч", "няв" тощо');
disp('Коли ви почуєте, наприклад, "Запишіть гуч тепер",');
disp('наберіть на клавіатурі гуч та натисніть Enter');

% Створення папки користувача для запису результатів:
if wt == 0
    %Тренування
    kat_results = ['ResultsTraining209_' lastname];% ім'я папки для рез-тів
else
    kat_results = ['Results209_' lastname]; % ім'я папки для рез-тів
end
mkdir(kat_results);    % створення основної папки для запису результатів

% Створ. папки для запису рез-тів досліджень
ngr = 4;
mkdir([kat_results '\Group' num2str(ngr)]);
addpath ('Programs', [kat_results '\Group' num2str(ngr)]);% шлях до програм

clc

%% 3. Запуск програми assistant_bin_3.m та одержання від неї вихідних даних:
truend = 1;
while (truend == 1)

[params] = assistant_bin_3(kat_results, NT,num_tables,num_SNRT,num_T60D);
N = params.N;    % номер еталонної таблиці звукосполучень
m = params.m;    % номер групи дослідів (в нас це номер 4)
suf = params.suf;

% 4. Завантаження wav-таблиці із потрібним номером
rfr1_tablename = ['rfr_table_' num2str(N) '_1.mat'];% ім'я 1-ї еталон.таблиці
rfr2_tablename = ['rfr_table_' num2str(N) '_2.mat'];% ім'я 2-ї еталон.таблиці
rfr3_tablename = ['rfr_table_' num2str(N) '_3.mat'];% ім'я 3-ї еталон.таблиці
wav_tablename = ['wav_table_' num2str(N) '.mat']; % ім'я wav-таблиці
wavs = load(wav_tablename);    % завантаження wav-таблиці
rslt_tablename = ['rslt_table_' suf '.mat'];    % ім'я таблиці результатів

%% 5. Прослуховування звукосполучення (за натиском довільної клавіші)

```

```

clc
disp([num2str(N) ' табл., ' num2str(SNRT(params.nsnr)) ' dB, ' num2str(T60D(params.nt60)) ' m]);
%disp('Якщо треба буде виправити попереднє звукосполучення - введіть 1'); % зміна Чайки
disp(' ');

% === початок спільної модернізації від О.Чайки та М.Ярошенко ===
symb=0;
k=1;
kk=[];
k1=randperm(nums);
while k <= nums
    if symb~= '1'
        inputSignal = wavs.(['w' num2str(k1(k))]);
        params.m = m;
        if (m == 2)||(m ==4)
            params.snr0 = SNRT(params.nsnr);    % таблиця значень потрібних SNR0
        end
        if k > 1 && k < nums
            %disp('Якщо треба виправити попередній набір - введіть 1'); % зміна Чайки
        end
        [outputSignal] = distortSystem_bin_50ms_cor(inputSignal, params);
        %pause;
        soundsc(outputSignal,params.Fs);
        textx=horzcat(num2str(k),' (' ,num2str(k1(k)),',',num2str(nums),') Введіть почуте
звукосполучення укр.мовою: ');
        symb = input(textx, 's');
        temprzlt.(['s' num2str(k1(k))]) = symb;
        k=k+1;
    elseif symb=='1'
        k=k-1;
        inputSignal = wavs.(['w' num2str(k1(k-1))]);
        params.m = m;
        if (m == 2)||(m ==4)
            params.snr0 = SNRT(params.nsnr);    % таблиця значень потрібних SNR0
        end
        [outputSignal] = distortSystem_bin_50ms_cor(inputSignal, params);
        %pause;
        soundsc(outputSignal,params.Fs);
        textx=horzcat(num2str(k-1),' (' ,num2str(k1(k-1)),',',num2str(nums),') Кореговане
звукосполучення: ');
        symb = input(textx, 's');
        temprzlt.(['s' num2str(k1(k-1))]) = symb;
    elseif size(symb,2)==0
        inputSignal = wavs.(['w' num2str(k1(k))]);
        params.m = m;
        if (m == 2)||(m ==4)
            params.snr0 = SNRT(params.nsnr);    % таблиця значень потрібних SNR0
        end
        if k > 1 && k < nums
            %disp('Якщо треба виправити попередній набір - введіть 1'); % зміна Чайки
        end
        [outputSignal] = distortSystem_bin_50ms_cor(inputSignal, params);
        %pause;
        soundsc(outputSignal,params.Fs);
        textx=horzcat(num2str(k),' (' ,num2str(k1(k)),',',num2str(nums),') Введіть почуте
звукосполучення: ');
        symb = input(textx, 's');
    end
end

```

```

    temprzlt(['s' num2str(k1(k))]) = symb;
    k=k+1;
end
if k==nums+1      % перевірка останнього звукосполучення
    perevirka = input('Правильно ввели останнє звукосполучення? Так - 1, Ні - 0 ');
    if perevirka==0
        k=k-1;
    end
end
end
end
% === кінець спільної модернізації від О.Чайки та М.Ярошенко ===

%% 6. Підрахунок частки правильно сприйнятих звукосполучень
rfrtbl_1 = load(rfr1_tablename); % завантаження 1-ї еталон.таблиці
rfrtbl_2 = load(rfr2_tablename); % завантаження 2-ї еталон.таблиці
rfrtbl_3 = load(rfr3_tablename); % завантаження 3-ї еталон.таблиці
num_right = 0; % початкове значення кількості правильних результатів
% clc
%disp('НАОЧНО СПІВСТАВТЕ РЕЗУЛЬТАТ ІЗ ЕТАЛОНАМИ');
%disp(' ');
%disp(['Еталон 1 ' ' Еталон 2 ' ' Еталон 3 ' ' Було введено']);
%disp(' ');
for k1 = 1:nums
    fldname = ['s' num2str(k1)];
    ref1 = rfrtbl_1.(fldname);
    ref2 = rfrtbl_2.(fldname);
    ref3 = rfrtbl_3.(fldname);
    rzlt = temprzlt.(fldname);
    % disp([ref1 ' ' ref2 ' ' ref3 ' ' rzlt]);
    if ((strcmp(rzlt, ref1))||strcmp(rzlt, ref2))||strcmp(rzlt, ref3))
        num_right = num_right + 1;
    end
end
corct = num_right/nums;
disp(' ');
% disp(['Правильно сприйнято: ' num2str(num_right) ' із ' num2str(nums)]);
disp(['Оцінка розбірливості мови становить: ' num2str(corct)]);

% запис відносної кількості правильних результатів в таблицю результатів та
% на диск
temprzlt('corct') = corct;
katname = [kat_results '\Group' num2str(m) '\ ' rzlt_tablename];
save (katname, '-struct', 'temprzlt');

if params.cntr4 == num_tables
    disp(' ');
    disp('Поздоровляємо із закінченням першої фази роботи!');
    disp(' ');
    disp('Одержані результати обробить програмою rzlts_processing_bin_209.m для побудови
графіків');
    disp('Графіки збережуть у форматі png');
    disp(' ');
    disp('Зархівуйте папку із результатами (включаючи графіки) та передайте її викладачу');
    return
end

% 7. Запит про подовження слухання таблиць

```



```
disp(' ');
truend = input('Подовжити досліді? Так - 1, Ні - 0 ');
end
```

Додаток 3

Програма rzlts_processing_bin_209.m

```
%% === rzlts_processing_bin_209.m === обробка результатів досліджень ===
%
% Мета - обробка результатів суб'єктивного оцінювання розбірливості мови
% (рез-ти - у полі "corct" кожної із таблиць, назва котрої починається із "rzlt_table_")
%
%          План програми:
%
% 1. Запит вхідних даних для програми assistant:
%   1.1. Прізвище виконавця
%   1.2. Три номери таблиць
%
% 2. Підрахунок частки правильно сприйнятих звукосполучень шляхом
% усереднення окремих результатів по 3-м таблицям
%
% В результаті одержуємо графіки залежностей розбірливості CVC(SNR) для
% різних значень T60D
%
% Copyright @ А.Продеус, вересень-жовтень 2017, серпень 2018, грудень 2020
% =====
%%
clc      % очистка командного окна
clear    % очистка пам'яті даних програми , т.н. workspace

% 1. Запит вхідних даних для програми assistant (прізвище, номери таблиць)
lastname = input('Вкажіть ваше прізвище латиницею (без лапок): ', 's');
disp('Вкажіть номери трьох таблиць, які використовувалися :');
NT = zeros(1,3);
NT(1) = input('Вкажіть номер першої таблиці: N1 = ');
NT(2) = input('Вкажіть номер другої таблиці: N2 = ');
NT(3) = input('Вкажіть номер третьої таблиці: N3 = ');
SNRT = [-6 0]; % таблиця значень потрібних SNR0
num_SNRT = length(SNRT);
T60D = [3.5 5.2 8.6 10.3]; % дистанції гучномовець-мікрофон
num_T60D = length(T60D);

disp(' ');
wt = input('Оберіть режим роботи? 1 - РОБОТА; 0 - ТРЕНУВАННЯ: '); %!!!
if wt == 0
    %numtren = input('Тренування: задайте кількість звукосполучень (від 3 до 5): ');
    nums = 3; % кількість звукосполучень в режимі ТРЕНУВАННЯ
else
    nums = 50; % кількість звукосполучень в режимі РОБОТА
end

%% Створення шляхів до папок із записаними результатами:
if wt == 0
    kat_results = ['RezultsTraining209_' lastname]; % ім'я папки для рез-тів Тренування
else
    kat_results = ['Rezults209_' lastname]; % ім'я папки для рез-тів Роботи
```

```

end
ngr = 4;
addpath ([kat_rezults '\Group' num2str(ngr)]);
clc

%% обробка результатів

val4 = zeros(num_T60D,num_SNRT,3); % виділення зони для окремих результатів
nclr = 1; % білий шум
for ntb = 1:3 % 3 таблиці
    for nsnr = 1:num_SNRT % кількість SNR0
        for nt60 = 1:num_T60D % кількість імпульсних характеристик
            rzltname = [kat_rezults '\Group' num2str(ngr) '\rzt_table_'...
                num2str(NT(ntb)) '_t' num2str(nt60) '_c' num2str(nclr)...
                '_s' num2str(nsnr) '.mat'];
            rzlt4 = load(rzltname);
            val4(nt60,nsnr,ntb) = rzlt4.('corct');
        end
    end
end
end
rzlt_gr4 = mean(val4,3); % усереднення результатів по трьом таблицям

%% побудова графіків результатів
% залежність CVC intellig від SNR
figure;
plot(SNRT,rzlt_gr4(1,:)); hold on;
plot(SNRT,rzlt_gr4(2,:),'-k'); hold on;
plot(SNRT,rzlt_gr4(3,:),'-r'); hold on;
plot(SNRT,rzlt_gr4(4,:),'s-b'); hold on;
%plot(SNRT,rzlt_gr4(5,:),'o-k'); hold on;
%plot(SNRT,rzlt_gr4(6,:),'--m'); hold on;
grid on;
xlabel('SNR (dB)'); ylabel('CVC intellig');
legend([num2str(T60D(1)) ' m'], [num2str(T60D(2)) ' m'],...
    [num2str(T60D(3)) ' m'], [num2str(T60D(4)) ' m']);

% залежність CVC intellig від T60D
figure;
plot(T60D,rzlt_gr4(:,1)); hold on;
plot(T60D,rzlt_gr4(:,2),'-k'); hold on;
grid on;
xlabel('Distance (m)'); ylabel('CVC intellig');
legend([num2str(SNRT(1)) ' dB'], [num2str(SNRT(2)) ' dB']);

%% збереження результатів усереднення в файлі-структурі
temptab.('gr4_rzlt_T60_1') = rzlt_gr4(1,:);
temptab.('gr4_rzlt_T60_2') = rzlt_gr4(2,:);
temptab.('gr4_rzlt_T60_3') = rzlt_gr4(3,:);
temptab.('gr4_rzlt_T60_4') = rzlt_gr4(4,:);
%temptab.('gr4_rzlt_T60_5') = rzlt_gr4(5,:);
%temptab.('gr4_rzlt_T60_6') = rzlt_gr4(6,:);
temptab.('gr4_SNR0') = SNRT;
temptab.('gr4_T60D') = T60D;

save ([kat_rezults '\mean_rezults.mat' ], '-struct', 'temptab');

```

Програма assistant_bin_3.m

```

function [params] = assistant_bin_3(kat_results, NT,num_tables,num_SNRT,num_T60D)
% Програма-асистент дає можливість переривати роботу в будь-який час
%
% Для запису результатів використовуються 2 папки:
% 1. Основна папка під назвою "Results_Surname", де Surname -
% прізвище студента
% 2. Папка під назвою "Group4", вложена в основну папку, для збереження результатів
% дослідів
%
% Copyright @ А.Продеус, верес 2017, серп 2018, лют 2019, груд 2020
% =====
ngr = 4;
maxnfiles4 = num_tables;
% формування списку імен файлів у папці Group4
pap_name_r = [kat_results '\Group' num2str(ngr)]; % шлях до папки
y1 = dir (pap_name_r); % список файлів в папці + 2 лишніх (перших)
z1 = y1(3:end); % список потрібних файлів в папці
kol_file = size(z1,1); % кількість файлів в папці

cntr = 0; % лічильник файлів із результатами
for nsnr = 1:num_SNRT % відношення сигнал-шум
    for nclr = 1:1 % білий шум
        for nt60 = 1:num_T60D % перебір імпульсних характеристик
            for ntb = 1:3 % номер таблиці
                cntr = cntr + 1;
                if (cntr > kol_file) && (cntr <= maxnfiles4)
                    params.suf = [num2str(NT(ntb)) '_t' num2str(nt60)...
                        '_c' num2str(nclr) '_s' num2str(nsnr)];
                    params.N = NT(ntb); params.m = ngr;
                    params.fs = 22050; % частота дискр-ції мовного сигналу
                    params.Fs = 44100; % частота дискр-ції ІХ
                    % завантаження таблиць БІНАУРАЛЬНИХ ІХ кімнат
                    rirs_LR = load('rirs209_table_LR.mat');
                    params.rir_LR = rirs_LR.(['r' num2str(nt60)]);
                    params.nt60 = nt60; % номер імпульсної хар-ки
                    params.nsnr = nsnr; % номер SNR
                    %noises = load('noises_table.mat'); % завант. табл. шумів
                    %params.noise = noises.(['n' num2str(nclr)]);
                    params.ncolor = nclr;
                    %params.cntr1 = maxnfiles1;
                    %params.cntr2 = maxnfiles2;
                    %params.cntr3 = maxnfiles3;
                    params.cntr4 = cntr;
                return
            end
        end
    end
end
disp(' ');
%disp('Закінчено групу дослідів');
end

```

Программа distortSystem_bin_50ms_cor.m

```
function [ outputSignal ] = distortSystem_bin_50ms_cor( inputSignal, params )
% === формирование речевого сигнала, искаженного шумом и реверберацией
% Це скоригована програма: в команді detVoice = vad_LP(inputSignalFs,Fs)
% замінено fs на Fs

% === ВХОДНЫЕ ДАННЫЕ: ===
% inputSignal - 1-канальный массив значений входного сигнала (речь с микрофона)
% fs - частота дискретизации входного сигнала
% Fs - частота дискретизации ИХ помещения
% params.snr0 - желаемое отношение сигнал-шум (в дБ)
% params.rir_LR - таблица бинауральных ИХ помещения (2-канальных ИХ)
% params.noise - шум (массив выборок)
%
% === ВЫХОДНЫЕ ДАННЫЕ: ===
% outputSignal - 2-канальный массив значений выходного сигнала
%
% Выходной сигнал следовало бы формировать как свертку исходного сигнала с
% ИХ помещения плюс шум, т.е.
%
% 
$$y(t) = x(t)*h(t)+n(t), \quad (1)$$

%
% где x(t) - входной сигнал; h(t) - ИХ помещения; n(t) - шум; y(t) -
% выходной сигнал. К сожалению, в нашем распоряжении имеются лишь
% нормированные по энергии ИХ помещения, что не позволяет строить
% зависимость разборчивости от отношения сигнал-шум:  $S = f(\text{SNR})$ , поскольку
% неизвестно, как сильно будет усиливать (или ослаблять) сигнал его
% свертка с ИХ.
%
% Поэтому поставим перед собой иной вопрос: насколько ИХ помещения хороша в
% качестве фильтра, подавляющего аддитивный шум.
%
% Такая постановка вопроса позволяет рассматривать иную модель:
%
% 
$$y(t) = [x(t)+n(t)]*h(t), \quad (2)$$

%
% где вполне уместно оценивать искомую зависимость  $S = f(\text{SNR})$  и
% сопоставлять ее с аналогичной зависимостью для модели
%
% 
$$y(t) = x(t)+n(t) \quad (3)$$

%
% Модель (2) реализуется поэтапно:
% 1) с помощью детектора пауз разделяются речевой сигнал и шум
% 2) stdSpeech и stdNoise вычисляются по выделенным участкам
% 3) оценивается snr:
% 
$$\text{snr} = 20*\log_{10}(\text{stdSpeech}/\text{stdNoise});$$

% 4) по snr и заданному snr0 вычисляется коррект. коэфф-т k0 для шума:
% 
$$k0 = 10^{(0.05*(\text{snr} - \text{snr0}))};$$

% 5) формируется сумма сигнала и откорректированного шума:
% 
$$\text{sum} = \text{sig} + k0*\text{noise};$$

% 6) смесь sum сворачивается с h(t), однако перед сверткой от h(t)
% отсекается хвост после 50 мс, поэтому в качестве "фильтра" служат
% ранние отражения.
%
% Copyright @ А.Продеус, вересень 2017-грудень 2020
% =====
```

```

fs = params.fs; % частота дискр-ції мовного сигналу
Fs = params.Fs; % частота дискр-ції ІХ

inputSignalFs = resample(inputSignal,Fs,fs); % передискретизація мови з 22050 Гц в 44100 Гц
slength = size(inputSignalFs,1); % довжина вхідного сигналу

detVoice = vad_LP(inputSignalFs,Fs); % детектор речи (виправлено - fs замінено на Fs)
nullSilence = inputSignalFs.*detVoice; % обнулення пауз
sgnl = nonzeros(nullSilence); % сигнал
varSpeech = var(sgnl); % дисперсія сигналу

noise = randn(slength,1); % белый гауссовый шум
varNoise = var(noise); % дисперсия шума
snr = 10*log10(varSpeech/varNoise); % имеющееся snr
k0 = 10^(0.05*(snr - params.snr0)); % коррект.коэфф-т
outputSig = inputSignalFs + k0*noise; % аддитивная смесь

rir = params.rir_LR; % чтение БИНАУРАЛЬНОЙ ІХ
% === выделение первых 50 мс ===
n50 = round(0.05*Fs); % номер выборки 50 мс
rir50 = rir(1:n50,:); % высечение 50 мс от ІХ
outputSignal = fftfilt(rir50,outputSig); % свертка ІХ с аддитивной смесью;
outputSignal = outputSignal/max(abs(outputSignal))*0.99; % нормир.

end

```